

Decidability Results for Multi-objective Stochastic Games

Romain Brenguier and Vojtěch Forejt

Department of Computer Science, University of Oxford, UK

Abstract. We study stochastic two-player turn-based games in which the objective of one player is to ensure several infinite-horizon total reward objectives, while the other player attempts to spoil at least one of the objectives. The games have previously been shown not to be determined, and an approximation algorithm for computing a Pareto curve has been given. The major drawback of the existing algorithm is that it needs to compute Pareto curves for finite horizon objectives (for increasing length of the horizon), and the size of these Pareto curves can grow unboundedly, even when the infinite-horizon Pareto curve is small.

By adapting existing results, we first give an algorithm that computes the Pareto curve for determined games. Then, as the main result of the paper, we show that for the natural class of stopping games and when there are two reward objectives, the problem of deciding whether a player can ensure satisfaction of the objectives with given thresholds is decidable. The result relies on intricate and novel proof which shows that the Pareto curves contain only finitely many points.

As a consequence, we get that the two-objective discounted-reward problem for unrestricted class of stochastic games is decidable.

1 Introduction

Formal verification is an area of computer science which deals with establishing properties of systems by mathematical means. Many of the systems that need to be modelled and verified contain controllable decisions, which can be influenced by a user, and behaviour which is out of the user's control. The latter can be further split into events whose presence can be quantified, such as failure rate of components, and events which are considered to be completely adversarial, such as acts of an attacker who wants to break into the system.

Stochastic turn-based games are used as a modelling formalism for such systems [4]. Formally, a stochastic game comprises three kinds of states, owned by one of three players: **Player 1**, **Player 2**, and the stochastic player. In each state, one or more transitions to successor states are available. At the beginning of a play, a token is placed on a distinguished initial state, and the player who controls it picks a transition and the token is moved to the corresponding successor state. This is repeated ad infinitum and a path, comprising an infinite sequence of states, is obtained. **Player 1** and **Player 2** have a free choice of transitions, and

the recipe for picking them is called a strategy. The stochastic player is bound to pick each transition with a fixed probability that is associated with it.

The properties of systems are commonly expressed using rewards, where numbers corresponding to gains or losses are assigned to states of the system. The numbers along the infinite paths are then summed, giving the total reward of an infinite path, intuitively expressing the energy consumed or the profit made along a system's execution. Alternatively, the numbers can be summed with a discounting $\delta < 1$, giving discounted reward. It formalises the fact that immediate gains matter more than future gains, and it is particularly important in economics where money received early can be invested and yield interest.

Traditionally, the aim of one player is to make sure the expected (discounted) total reward exceeds a given bound, while the other player tries to ensure the opposite. We study the *multi-objective problem* in which each state is given a tuple of numbers, for example corresponding to both the profit made on visiting the state, and the energy spent. Subsequently, we give a bound on both profit and energy, and **Player 1** attempts to ensure that the expected total profit and expected total energy exceed (or do not exceed) the given bound, while **Player 2** tries to spoil this by making sure that at least one of the goals is not met.

The problem has been studied in [7], where it has been shown that Pareto optimal strategies might not exist, and the game might not be determined (for some bounds neither of the players have ε -optimal strategies). A value iteration algorithm has been given for approximating the Pareto curve of the game, i.e. the bounds **Player 1** can ensure. The algorithm successively computes, for increasing n , the sets of bounds **Player 1** can ensure if the length of the game is restricted to n steps. The approach has two major drawbacks. Firstly, the algorithm cannot decide, for given bounds, if **Player 1** can achieve them. Secondly, it does not scale well since the representation of the sets can grow with increasing n , even if the ultimate Pareto curve is small.

The above limitations show that it is necessary to design alternative solution approaches. One of the promising directions is to characterise the shape of the set of achievable bounds, for computing it efficiently. The value iteration of [7] allows us to show that the sets are convex, but no further observations can be made, in particular it is not clear whether the sets are convex polyhedra, or if they can have infinitely many extremal points. The main result of our paper shows that for two-objective case and stopping games, the sets are indeed convex polyhedra, which directly leads to a decision algorithm. We believe that our proof technique is of interest on its own. It proceeds by assuming that there is an accumulation point on the Pareto curve, and then establishes that there must be an accumulation point in one of the successor states such that the *slope* of the Pareto curves in the accumulation points are equal. This allows us to obtain a cycle in the graph of the game in which we can “follow” the accumulation points and eventually revisit some of them infinitely many times. By further analysing slopes of points on the Pareto curves that are close to the accumulation point, we show that there are two points on the curve that are sufficiently far from each other yet have the same slope, which contradicts the assumption that they are near an accumulation point.

Our results also yield novel important contributions for non-stochastic games. Although there have recently been several works on non-stochastic games with multiple objectives, they a priori restrict to deterministic strategies, by which the associated problems become fundamentally different. It is easy to show that enabling randomisation of strategies extends the bounds **Player 1** can achieve, and indeed, even in other areas of game-theory randomised strategies have been studied for decades: the fundamental theorem of game theory is that every finite game admits a *randomised* Nash equilibrium [13].

Related work. In the area of stochastic games, single-objective problems are well studied. For reachability objectives the games are determined and the problem of existence of an optimal strategy achieving a given value is in $\text{NP} \cap \text{co-NP}$ [8]; same holds for total reward objectives. In the multi-objective setting, [7] gives a value iteration algorithm for the multi-objective total reward problem. Although value iteration converges to the correct result, it does so only in infinite number of steps. It is further shown in [7] that when **Player 1** is restricted to only use deterministic strategies, the problem becomes undecidable; the proof relies fundamentally on the strategies being deterministic and it is not clear how it can be extended to randomised strategies. The work of [1] extends the equations of [7] to expected energy objectives, and mainly concerns itself with a variant of multi-objective mean-payoff reward, where the objective is a “satisfaction objective” requiring that there is a set of runs of a given probability on which all mean payoff rewards exceed a given bound (i.e., expected values are not considered). [1] only studies existence of finite-memory strategies and the probability bound 1; this restriction has very recently been lifted by [3], which shows that even unrestricted satisfaction objective problem is coNP -complete.

In non-stochastic games, multi-objective optimisation has been studied for multiple mean-payoff objectives and energy games [16]. A comprehensive analysis of the complexity of synthesis of optimal strategies has been given [5], and it has been shown that a variant of the problem is undecidable [15]. The work of [2] studies the complexity of problems related to exact computation of Pareto curves for multiple mean-payoff objectives. In [11], interval objectives are studied for total, mean-payoff and discounted reward payoff functions. The problems for interval objectives are a special kind of multi-objective problems that require the payoff to be within a given interval, as opposed to the standard single-objective setting where the goal is to exceed a given bound. As mentioned earlier, all the above works for non-stochastic games a priori restrict the players to use deterministic strategies, and hence the problems exhibit completely different properties than the problem we study.

Our contribution. We give the following novel decidability results. Firstly, we show that the problem for *determined* stochastic games is decidable. Then, as the main result of the paper, we show that for non-determined games which also satisfy the stopping assumption and for two objectives, the set of achievable bounds forms a convex polyhedron. This immediately leads to an algorithm for computing Pareto curves, and we obtain the following novel results as corollaries.

- Two-objective discounted-reward problem for stochastic games is decidable.
- Two-objective total-reward problem for stochastic stopping games is decidable.

Although we phrase our results in terms of stochastic games, to our best knowledge, the above results also yield novel decidability results for multi-objective *non-stochastic games* when randomisation of strategies is allowed.

Outline of the paper. In Sec. 3, we show a simple algorithm that works for determined games and show how to decide whether a stopping game is determined. In Sec. 4, we give decidability results for two-objective stopping games.

2 Preliminaries on stochastic games

We begin this section by introducing the notation used throughout the paper. Given a vector $\mathbf{v} \in \mathbb{R}^n$, we use v_i to refer to its i -th component, where $1 \leq i \leq n$. The comparison operator \leq on vectors is defined to be the componentwise ordering: $\mathbf{u} \leq \mathbf{v} \Leftrightarrow \forall i \in [1, n]. \mathbf{u}_i \leq \mathbf{v}_i$. We write $\mathbf{u} < \mathbf{v}$ when $\mathbf{u} \leq \mathbf{v}$ and $\mathbf{u} \neq \mathbf{v}$. Given two vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$, the *dot product* of \mathbf{u} and \mathbf{v} is defined by $\mathbf{u} \cdot \mathbf{v} = \sum_{i=1}^n \mathbf{u}_i \cdot \mathbf{v}_i$.

The sum of two sets of vectors $U, V \subseteq \mathbb{R}^n$ is defined by $U + V = \{\mathbf{u} + \mathbf{v} \mid \mathbf{u} \in U, \mathbf{v} \in V\}$. Given a set $V \subseteq \mathbb{R}^n$, we define the *downward closure* of V as $\text{dwc}(V) \stackrel{\text{def}}{=} \{\mathbf{u} \mid \exists \mathbf{v} \in V. \mathbf{u} \leq \mathbf{v}\}$, and we use $\text{conv}(V)$ for the *convex closure* of V , i.e. the set of all \mathbf{v} for which there are $\mathbf{v}^1, \dots, \mathbf{v}^n \in V$ and $w_1 \dots w_n \in [0, 1]$ such that $\sum_{i=1}^n w_i = 1$ and $\mathbf{u} = \sum_{i=1}^n w_i \cdot \mathbf{v}^i$. An *extremal point* of a set $X \subseteq \mathbb{R}^n$ is a vector $x \in X$ that is not a convex combination of other points in X , i.e. $x \notin \text{conv}(X \setminus \{x\})$.

A function $f: \mathbb{R} \rightarrow \mathbb{R}$ is concave whenever for all $x, y \in \mathbb{R}$ and $t \in [0, 1]$ we have $f(t \cdot x + (1-t) \cdot y) \geq t \cdot f(x) + (1-t) \cdot f(y)$. Given $x \in \mathbb{R}$, the *left slope* of f in x is defined by $\text{lslope}(f, x) \stackrel{\text{def}}{=} \lim_{x' \rightarrow x^-} \frac{f(x) - f(x')}{x - x'}$. Similarly the *right slope* is defined by $\lim_{x' \rightarrow x^+} \frac{f(x) - f(x')}{x - x'}$. Note that if f is concave then both limits are well-defined, because by concavity $\frac{f(x) - f(x')}{x - x'}$ is monotonic in x' ; nevertheless, the left and right slope might still not be equal.

A point $\mathbf{p} \in \mathbb{R}^2$ is an *accumulation point* of f if $f(\mathbf{p}_1) = \mathbf{p}_2$ and for all $\varepsilon > 0$, there exists $x \neq \mathbf{p}_1$ such that $(x, f(x))$ is an extremal point of f and $|\mathbf{p}_1 - x| < \varepsilon$. Moreover, \mathbf{p} is a *left (right) accumulation point* if in the above we in addition have $x < \mathbf{p}_1$ (resp. $x > \mathbf{p}_1$). We sometimes slightly abuse notation by saying that x is an extremal point when $(x, f(x))$ is an extremal point, and similarly for accumulation points.

A *discrete probability distribution* (or just *distribution*) over a (countable) set S is a function $\mu: S \rightarrow [0, 1]$ such that $\sum_{s \in S} \mu(s) = 1$. We write $\mathcal{D}(S)$ for the set of all distributions over S , and use $\text{supp}(\mu) = \{s \in S \mid \mu(s) > 0\}$ for the *support set* of $\mu \in \mathcal{D}(S)$.

We now define turn-based stochastic two-player games together with the concepts of strategies and paths of the game. We then present the objectives that are studied in this paper and the associated decision problems.

Stochastic games. A *stochastic (two-player) game* is a tuple $\mathcal{G} = \langle S, (S_\square, S_\diamond, S_\circ), \Delta \rangle$ where S is a finite set of states partitioned into sets S_\square , S_\diamond , and S_\circ ; $\Delta : S \times S \rightarrow [0, 1]$ is a probabilistic transition function such that $\Delta(s, t) \in \{0, 1\}$ if $s \in S_\square \cup S_\diamond$ and $\sum_{t \in S} \Delta(s, t) = 1$ if $s \in S_\circ$.

S_\square and S_\diamond represent the sets of states controlled by Player 1 and Player 2, respectively, while S_\circ is the set of stochastic states. For a state $s \in S$, the set of successor states is denoted by $\Delta(s) \stackrel{\text{def}}{=} \{t \in S \mid \Delta(s, t) > 0\}$. We assume that $\Delta(s) \neq \emptyset$ for all $s \in S$. A state from which no other states except for itself are reachable is called *terminal*, and the set of terminal states is denoted by $\text{Term} \stackrel{\text{def}}{=} \{s \in S \mid \Delta(s) = \{s\}\}$.

Paths. An *infinite path* λ of a stochastic game \mathcal{G} is a sequence $(s_i)_{i \in \mathbb{N}}$ of states such that $s_{i+1} \in \Delta(s_i)$ for all $i \geq 0$. A *finite path* is a prefix of such a sequence. For a finite or infinite path λ we write $\text{len}(\lambda)$ for the number of states in the path. For $i < \text{len}(\lambda)$ we write λ_i to refer to the i -th state s_{i-1} of $\lambda = s_0 s_1 \dots$ and $\lambda_{\leq i}$ for the prefix of λ of length $i + 1$. For a finite path λ we write $\text{last}(\lambda)$ for the last state of the path. For a game \mathcal{G} we write $\Omega_{\mathcal{G}}^+$ for the set of all finite paths, and $\Omega_{\mathcal{G}}$ for the set of all infinite paths, and $\Omega_{\mathcal{G}, s}$ for the set of infinite paths starting in state s . We denote the set of paths that reach a state in $T \subseteq S$ by $\Diamond T \stackrel{\text{def}}{=} \{\lambda \in \Omega_{\mathcal{G}} \mid \exists i. \lambda_i \in T\}$.

Strategies. We write $\Omega_{\mathcal{G}}^\square$ and $\Omega_{\mathcal{G}}^\diamond$ for the finite paths that end with a state of S_\square and S_\diamond , respectively. A *strategy* of Player 1 is a function $\pi : \Omega_{\mathcal{G}}^\square \rightarrow \mathcal{D}(S)$ such that $s \in \text{supp}(\pi(\lambda))$ only if $\Delta(\text{last}(\lambda), s) = 1$. We say that π is *memoryless* if $\text{last}(\lambda) = \text{last}(\lambda')$ implies $\pi(\lambda) = \pi(\lambda')$, and *deterministic* if $\pi(\lambda)$ is Dirac for all $\lambda \in \Omega_{\mathcal{G}}^+$, i.e. $\pi(\lambda)(s) = 1$ for some $s \in S$. A strategy σ for Player 2 is defined similarly replacing $\Omega_{\mathcal{G}}^\square$ with $\Omega_{\mathcal{G}}^\diamond$. We denote by Π and Σ the sets of all strategies for Player 1 and Player 2, respectively.

Probability measures. A stochastic game \mathcal{G} , together with a strategy pair $(\pi, \sigma) \in \Pi \times \Sigma$ and an initial state s , induces an infinite Markov chain on the game (see e.g. [6]). We denote the probability measure of this Markov chain by $\mathbb{P}_{\mathcal{G}, s}^{\pi, \sigma}$. The expected value of a measurable function $g : S^\omega \rightarrow \mathbb{R}_{\pm\infty}$ is defined as $\mathbb{E}_{\mathcal{G}, s}^{\pi, \sigma}[g] \stackrel{\text{def}}{=} \int_{\Omega_{\mathcal{G}, s}} g d\mathbb{P}_{\mathcal{G}, s}^{\pi, \sigma}$. We say that a game \mathcal{G} is a *stopping game* if, for every strategy pair (π, σ) , a terminal state is reached with probability 1, i.e. $\mathbb{P}_{\mathcal{G}, s}^{\pi, \sigma}(\Diamond \text{Term}) = 1$ for all s .

Total reward. A reward function $\varrho : S \rightarrow \mathbb{Q}$ assigns a reward to each state of the game. We assume the rewards are 0 in all terminal states. The *total reward* of a path λ is $\varrho(\lambda) \stackrel{\text{def}}{=} \sum_{j \geq 0} \varrho(\lambda_j)$. Given a game \mathcal{G} , an initial state s , a vector of n rewards $\mathbf{\varrho}$ and a vector of n bounds $\mathbf{z} \in \mathbb{R}^n$, we say that a pair of strategies (π, σ) *yields* an objective $\text{totrew}(\mathbf{\varrho}, \mathbf{z})$ if $\mathbb{E}_{\mathcal{G}, s}^{\pi, \sigma}[\mathbf{\varrho}_i] \geq \mathbf{z}_i$ for all $1 \leq i \leq n$. A strategy $\pi \in \Pi$ *achieves* $\text{totrew}(\mathbf{\varrho}, \mathbf{z})$ if for all σ we have that (π, σ) yields $\text{totrew}(\mathbf{\varrho}, \mathbf{z})$; the vector \mathbf{z} is then called *achievable*, and we use \mathcal{A}_s for the set of all achievable vectors. A strategy $\sigma \in \Sigma$ *spoils* $\text{totrew}(\mathbf{\varrho}, \mathbf{z})$ if for no $\pi \in \Pi$, the tuple (π, σ) yields $\text{totrew}(\mathbf{\varrho}, \mathbf{z})$. Note that lower bounds (objectives $\mathbb{E}_{\mathcal{G}, s}^{\pi, \sigma}[\mathbf{\varrho}_i] \leq \mathbf{z}_i$) can be modelled by upper bounds after multiplying all rewards and bounds by -1 .

A (lower) Pareto curve in s is the set of all maximal \mathbf{z} such that for all $\varepsilon > 0$ there is $\pi \in \Pi$ that achieves the objective $\text{totrew}(\boldsymbol{\varrho}, \mathbf{z} - \varepsilon)$. We use f_s for the Pareto curve, and for the two-objective case we treat it as a function, writing $f_s(x) = y$ when $(x, y) \in f_s$. We say that a game is *determined* if for all states, every bound can be spoiled or lies in the downward closure of the Pareto curve¹. Note that the downward closure of the Pareto curve equals the closure of \mathcal{A}_s .

Discounted reward. Discounted games play an important role in game theory. In these games, the rewards have a discount factor $\delta \in (0, 1)$ meaning that the reward received after j steps is multiplied by δ^j , and so a discounted reward of a path λ is then $\varrho(\lambda, \delta) = \sum_{j \geq 0} \varrho(\lambda_j) \cdot \delta^j$. We define the notions of achieving, spoiling and Pareto curves for discounted reward $\text{disrew}(\boldsymbol{\varrho}, \delta, \mathbf{z})$ in the same way as for total reward. Since the problems for discounted reward can easily be encoded using the total reward framework (by adding before each state a stochastic state from which with probability $(1 - \delta)$ we transition to a terminal state), from now on we will concentrate on total reward, unless specified otherwise.

The problems. In this paper we study the following decision problems.

Definition 1 (Total-reward problem). *Given a stochastic game \mathcal{G} , an initial state s_0 , and vectors of reward functions $\boldsymbol{\varrho}$ and thresholds \mathbf{z} , is $\text{totrew}(\boldsymbol{\varrho}, \mathbf{z})$ achievable from s_0 ?*

Definition 2 (Discounted-reward problem). *Given a stochastic game \mathcal{G} , an initial state s_0 , vectors of reward functions $\boldsymbol{\varrho}$ and thresholds \mathbf{z} , and a discount factor $\delta \in (0, 1)$, is $\text{disrew}(\boldsymbol{\varrho}, \delta, \mathbf{z})$ achievable from s_0 ?*

In the particular case when n above is 2, we speak about *two-objective* problems.

Simplifying assumption. In order to keep the proofs simple, we will assume that each non-terminal state has exactly two successors and that only the states controlled by Player 2 have weights different from 0. Note that any stochastic game can be transformed into an equivalent game with this property in polynomial time, so we do not lose generality by this assumption.

Example 3 (Floor heating problem). As an example illustrating the definitions, as well as possible applications of our results, we consider a simplified version of the smart-house case study presented in [12] with a difference that we model both user comfort and energy consumption. Player 1, representing a controller, decides which rooms are heated, while the Player 2 represents the configuration of the house, for instance which door and windows are open, which cannot be influenced by the controller. The temperature in another room changes based on additional probabilistic factors. We illustrate this example in Fig. 1 and a simple model as a stochastic game is given in Fig. 2 (left). We have to control

¹ The reader might notice that in some works, games are said to be determined when each vector can be either achieved by one player, or spoiled by the other. This is not the case of our definition, where the notion of determinacy is *weaker* and only requires ability to spoil or achieve up to arbitrarily small ε .

the floor heating of two rooms in a house, by opening at most one of the valves V_1 and V_2 at a time.

The state of each room is either cold or hot, for instance in state H, C , the first room is warm while the second one is cold, and the third room has unknown temperature. Weights on the first dimension represent the energy consumption of the system while the second represent the comfort inside the house. **Player 2**

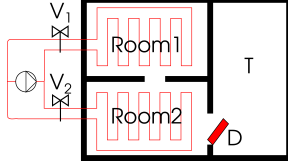


Fig. 1: A house with controllable floor heating in two rooms.

chooses whether the door is opened or not and stochastic states determine the contribution of the other rooms: for instance from (H, C) if the second player chooses that the door is opened then depending on whether the temperature of the other room is low or high, room 2 can either stay cold or get heated through the door, and the next state in that case is (H, H) which is the terminal state. The objective is to optimise energy consumption and comfort until both rooms are warm. The Pareto curve for a few states of the game is depicted in Fig. 2 (right).

2.1 Equations for lower value

We recall the results of [7,1] showing that for stopping games the sets of achievable points \mathcal{A}_s are the unique solution to the sets of equations defined as follows:

$$X_s = \begin{cases} \text{dwc}(\{(0, \dots, 0)\}) & \text{if } s \in \text{Term} \\ \text{dwc}(\text{conv}(\bigcup_{t \in \Delta(s)} X_t)) & \text{if } s \in S_{\square} \\ \underline{g}(s) + \text{dwc}(\bigcap_{t \in \Delta(s)} X_t) & \text{if } s \in S_{\diamond} \\ \text{dwc}(\sum_{t \in \Delta(s)} \Delta(s, t) \cdot X_t) & \text{if } s \in S_{\circ} \end{cases}$$

The equations can be used to design a value-iteration algorithm that iteratively computes sets X_s^i for increasing i : As a base step we have $X_s^0 = \text{dwc}(\mathbf{0})$ (where $\mathbf{0} = (0, \dots, 0)$); we then substitute X_s^i for X_s on the right-hand side of the equations, and obtain X_s^{i+1} as X_s on the left-hand side. The sets X_s^i so obtained converge to the least fixpoint of the equations above [7,1]. As we will show later, the sets X_s^i might be getting increasingly complex even though the actual solution X_s only comprises two extremal points.

3 Determined games

In this section we present a simple algorithm which works under the assumption that the game is determined. For stopping games, we then give a procedure to decide whether a game is determined.

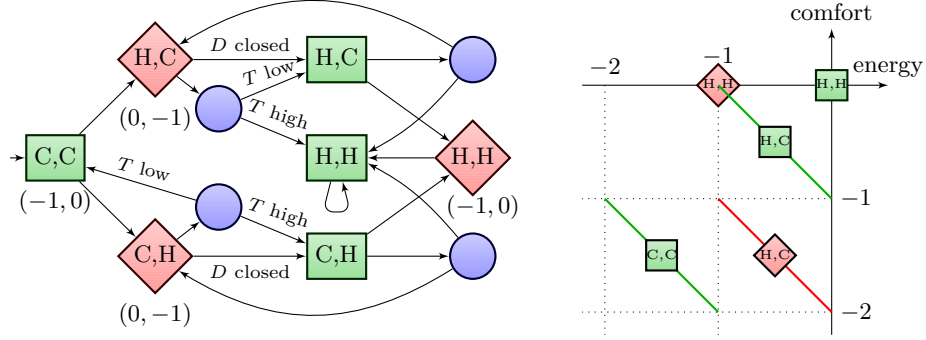


Fig. 2: A stochastic two-player game modelling the floor heating problem. Vectors under states denote a reward function when it is not $(0, 0)$. All probabilistic transitions have probability $\frac{1}{2}$. Pareto curves of a few states of the game are depicted on the right.

Theorem 3. *There is an algorithm working in exponential time, which given a determined stochastic two-player game, computes its Pareto-curve.*

For the proof of the theorem we will make use of the following:

Theorem 4 ([7, Thm. 7]). *Suppose Player 2 has a strategy σ such that for all π of Player 1 there is at least one $1 \leq i \leq n$ with $\mathbb{E}_{\mathcal{G},s}^{\pi,\sigma}(\mathbf{q}_i) < \mathbf{z}_i$. Then Player 2 has a memoryless deterministic strategy with the same properties.*

From the above theorem we obtain the following lemma.

Lemma 5. *The following two statements are equivalent for determined games:*

- *A given point \mathbf{z} lies in the downward closure of the Pareto curve for s .*
- *For all memoryless deterministic strategies σ of Player 2, there is a strategy π of Player 1 such that (π, σ) yield $\text{totrew}(\mathbf{q}, \mathbf{z})$.*

Thus, to compute the Pareto curve for a determined game \mathcal{G} , it is sufficient to consider all memoryless deterministic strategies $\sigma_1, \sigma_2, \dots, \sigma_m$ of Player 2 and use [9] to compute the Pareto curves $f_s^{\sigma_i}$ for the games \mathcal{G}^{σ_i} induced by \mathcal{G} and σ_i (i.e. \mathcal{G}^{σ_i} is obtained from \mathcal{G} by turning all $s \in S_\diamond$ to stochastic vertices and stipulating $\Delta(s, t) = \sigma_i(s)$ for all successors t of s ; in turn, \mathcal{G}^{σ_i} is a Markov decision process), and obtain the Pareto curve for \mathcal{G} as the pointwise minimum $V_s := \min_{1 \leq i \leq m} f_s^{\sigma_i}$.

To decide if a stopping game is determined, it is sufficient to take the downward closures of solutions V_s and check if they satisfy the equations from Sec. 2.1. Since in stopping games the solution of the equations is unique, if the sets are a solution they are also the Pareto curves and the game is determined. If any of the equations are not satisfied, then V_s are not the Pareto curves and the game is not determined. Note that for non-stopping games the above approach does

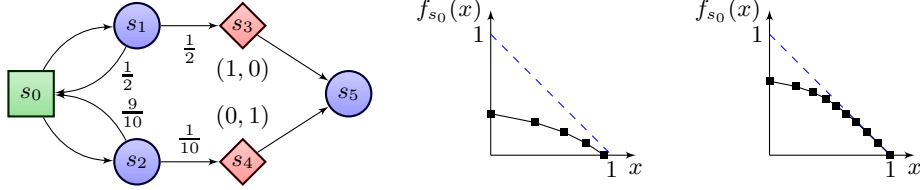


Fig. 3: An example showing that value iteration might produce Pareto curves with unboundedly many extremal points.

not work: even if the sets do not change by applying one step of value iteration, it is still possible that the solution is not the least fixpoint, and so we cannot infer any conclusion.

4 Games with two objectives

We start this section by showing that the existing value iteration algorithm presented in Sec. 2.1 might iteratively compute sets X_s^i with increasing number of extremal points, although the actual resulting set X_s (and the associated Pareto curve f_s) is very simple. Consider the game from Fig. 3 (left). Applying the value-iteration algorithm given by the equations from Sec. 2.1 for n steps gives a Pareto curve in s_0 with $n - 1$ extremal points. Each extremal point corresponds to a strategy π_i that in s_0 chooses to go to s_2 when the number of visits of s_0 is less than i , and after that chooses to go to s_1 . The upper bounds of the sets X_s^n for $n = 5$ and $n = 10$ are drawn in Fig. 3 (centre and right, respectively) using solid line, and their extremal points are marked with dots. The Pareto curve f_s is drawn with dashed blue line, and it consists of two extremal points, $(0, 1)$ and $(1, 0)$.

We now proceed with the main result of this section, the decidability of the two-objective strategy synthesis problem for stopping games. The result can be obtained from the following theorem.

Theorem 6. *If \mathcal{G} is a stopping stochastic two-player game with two objectives, and s a state of \mathcal{G} then the Pareto curve f_s has only finitely many extremal points.*

The above theorem can be used to design the following algorithm. For a fixed number k , we create a formula φ_k over $(\mathbb{R}, +, \cdot, \leq)$ which is true if and only if for each $s \in S$ there are points $\mathbf{p}^{s,1}, \dots, \mathbf{p}^{s,k}$ such that the sets $V_s \stackrel{\text{def}}{=} \text{dwc}(\text{conv}(\{\mathbf{p}^{s,1}, \dots, \mathbf{p}^{s,k}\}))$ satisfy the equations from Sec. 2.1. Using [14] we can then successively check validity of φ_k for increasing k , and Thm. 6 guarantees that we will eventually get a formula which is valid, and it immediately gives us the Pareto curve. We get the following result as a corollary.

Corollary 7. *Two-objective total reward problem is decidable for stopping stochastic games, and two-objective discounted-reward problem is decidable for stochastic games.*

Outline of the proof of Thm. 6. The proof of Thm. 6 proceeds by assuming that there are infinitely many extremal points on the Pareto curve, and then deriving a contradiction. Firstly, because the game is stopping, an upper bound on the expected total reward that can be obtained with respect to a single total reward objective is $M := \sum_{i=0}^{\infty} (1 - p_{\min}^{|S|}) \cdot \varrho_{\max}^{|S|}$ where $p_{\min} = \min\{\Delta(s, s') \mid \Delta(s, s') > 0\}$ is the smallest transition probability, and $\varrho_{\max} = \max_{i \in \{1, 2\}} \max_{s \in S} \varrho_i(s)$ is the maximal reward assigned to a state. Thus, the Pareto curve is contained in a compact set, and this implies that there is an accumulation point on it. In Sec. 4.1, we show that we can follow one accumulation point \mathbf{p} from one state to one of its successors, while preserving the same left slope. Moreover, in the neighbourhood of the accumulation point the rate at which the right slope decreases is quite similar to the decrease in the successors, in a way that is made precise in Lem. 9, 10, and 11. This is with the exception of some stochastic states for which the decrease strictly slows down when going to the successors: we will exploit this fact to get a contradiction. We construct a transition system $T_{s_0, \mathbf{p}}$, which keeps all the paths obtained by following the accumulation point \mathbf{p} from s_0 . We show that if \mathcal{G} is a stopping game, then we can obtain a path in $T_{s_0, \mathbf{p}}$ which visits stochastic states for which the decrease of the right slope strictly slows down. This relies on results for *inverse betting games*, which are presented in Sec. 4.2. Since this decrease can be repeated and there are only finitely many reachable states in $T_{s_0, \mathbf{p}}$, we show in Sec. 4.3 that the decrease of the right slope must be zero somewhere, meaning that the curve is constant in the neighbourhood of an accumulation point, which is a contradiction.

We will rely on the properties of the equations from Sec. 2.1 and the left and right slopes of the Pareto curve. Note that we introduced the notion of slope only for two-dimensional sets, and so our proofs only work for two dimensions. Generalisations of the concept of slopes exist for higher dimensions, but simple generalisation of our lemmas would not be valid, as we will show later. Hence, in the remainder of this section, we focus on the two-objective case. For the simplicity of presentation, we will present all claims and proofs for *left* accumulation points. The case of right accumulation points is analogous.

4.1 Mapping accumulation points to successor states

We start by enumerating some basic but useful properties of the Pareto curve and its slopes. First notice that it is a continuous concave function and we can prove the following:

Lemma 8. *Let f be a continuous concave function defined on $[a, b]$.*

1. *If $a < x < x' \leq b$ are two reals for which $lslope(f)$ is defined, then $lslope(f, x) \geq rslope(f, x) \geq lslope(f, x')$.*
2. *If (x, x') contains an extremal point of f then $lslope(f, x) \neq lslope(f, x')$.*
3. *If $x \in (a, b]$, then $\lim_{x' \rightarrow x^-} lslope(f, x') = \lim_{x' \rightarrow x^-} rslope(f, x') = lslope(f, x)$.*

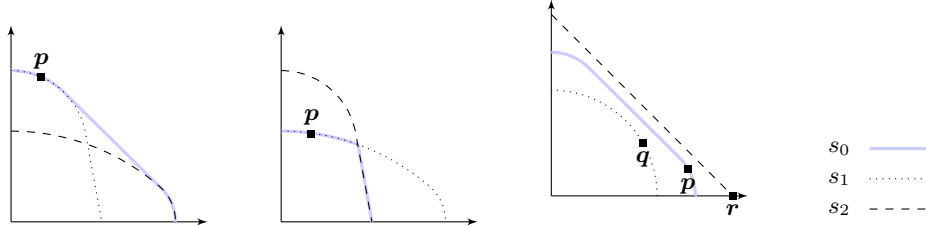


Fig. 4: An example of Pareto curve in a state s_0 with two successors s_1 and s_2 , for the case of $s_0 \in S_{\square}$ (left), $s_0 \in S_{\diamond}$ (centre), and $s_0 \in S_{\circ}$ with uniform probabilities on transitions (right). In each case, the curve in s_0 has infinitely many accumulation points.

To prove Thm. 6, we will use the equations from Sec. 2.1 to describe how accumulation points on a Pareto curve for s “map” to accumulation points on successors.

Lemma 9. *Let s_0 be a Player 1 state with two successors s_1 and s_2 , and let \mathbf{p} be a left accumulation point of f_{s_0} . Then there is $\eta(s_0, \mathbf{p}) > 0$ such that for all $\varepsilon \in (0, \eta(s_0, \mathbf{p}))$, there is $s' \in \{s_1, s_2\}$ such that: 1. \mathbf{p} is a left accumulation point in $f_{s'}$; 2. $\text{lslope}(s_0, \mathbf{p}_1) = \text{lslope}(s', \mathbf{p}_1)$; 3. $f_{s_0}(\mathbf{p}_1 - \varepsilon) \geq f_{s'}(\mathbf{p}_1 - \varepsilon)$ and $\text{rslope}(f_{s_0}, \mathbf{p}_1 - \varepsilon) \geq \text{rslope}(f_{s'}, \mathbf{p}_1 - \varepsilon)$.*

Proof (Sketch). The point 1. follows from the fact that every extremal point in the Pareto curve for s_0 must be an extremal point in one of the successors. This is illustrated in Fig. 4 (left): \mathbf{p} which is an extremal point for s_0 is also an extremal point for s_1 . The point 2. follows because from a sequence of extremal points $(\mathbf{p}^i)_{i \geq 0}$ on the Pareto curve of s_0 that converge to \mathbf{p} , we can select a subsequence that gives extremal points on s' that converge to the left accumulation point \mathbf{p} on s' . Finally, to prove 3. we use the fact that the right slope of f_{s_0} is always between those of f_{s_1} and of f_{s_2} . \square

Lemma 10. *Let s_0 be a Player 2 state with two successors s_1 and s_2 , and let \mathbf{p} be a left accumulation point of f_{s_0} . There is $\eta(s_0, \mathbf{p}) > 0$ such that for all $\varepsilon \in (0, \eta(s_0, \mathbf{p}))$, there is $s' \in \{s_1, s_2\}$, such that: 1. $\mathbf{p} - \boldsymbol{\varrho}(s_0)$ is a left accumulation point in $f_{s'}$; 2. $\text{lslope}(s_0, \mathbf{p}_1) = \text{lslope}(s', \mathbf{p}_1 - \boldsymbol{\varrho}_1(s_0))$; 3. $f_{s_0}(\mathbf{p}_1 - \varepsilon) = f_{s'}(\mathbf{p}_1 - \varepsilon - \boldsymbol{\varrho}_1(s_0))$ and $\text{rslope}(f_{s_0}, \mathbf{p}_1 - \varepsilon) = \text{rslope}(f_{s'}, \mathbf{p}_1 - \varepsilon - \boldsymbol{\varrho}_1(s_0))$.*

Proof (Sketch). A crucial observation here is that $f_{s_0}(\mathbf{p}_1^i)$ is either $\boldsymbol{\varrho}_2(s_0) + f_{s_1}(\mathbf{p}_1^i - \boldsymbol{\varrho}_1(s_0))$ or $\boldsymbol{\varrho}_2(s_0) + f_{s_2}(\mathbf{p}_1^i - \boldsymbol{\varrho}_1(s_0))$. This is illustrated in Fig. 4 (center): $f_{s_0}(\mathbf{p}_1) = \boldsymbol{\varrho}_2(s_0) + f_{s_1}(\mathbf{p}_1 - \boldsymbol{\varrho}_1(s_0))$ (there $\boldsymbol{\varrho}(s_0) = (0, 0)$). Hence when we take a sequence $(\mathbf{p}_1^i)_{i \in \mathbb{N}}$, for some $\ell \in \{1, 2\}$ the value $f_{s_0}(\mathbf{p}_1^i)$ equals $\boldsymbol{\varrho}_2(s_0) + f_{s_\ell}(\mathbf{p}_1^i - \boldsymbol{\varrho}_1(s_0))$ infinitely many times. From this we get a converging sequence of points in s_ℓ , and obtain that the left slopes are equal in s_0 and s_ℓ . By further arguing that in any left neighbourhood of $\mathbf{p}_1 - \boldsymbol{\varrho}_1(s_0)$ we can find infinitely many points with different left slopes, we obtain that there are also infinitely many extremal points in the neighbourhood and hence $\mathbf{p}_1 - \boldsymbol{\varrho}_1(s_0)$ is a left accumulation point.

As for the last item, the important observation here is that if at some point \mathbf{p}' , f_{s_1} is strictly below f_{s_2} then the right slope of f_{s_0} corresponds to that of f_{s_1} , and if f_{s_1} equals f_{s_2} then the right slope of f_{s_0} corresponds to the minimum of the right slopes of f_{s_1} and f_{s_2} (it is also interesting to note that the left slope corresponds to the maximum of the two). \square

Lemma 11. *Let s_0 be a stochastic state with two successors s_1 and s_2 , and \mathbf{p} a left accumulation point of f_{s_0} . There are points \mathbf{q} and \mathbf{r} on f_{s_1} and f_{s_2} respectively such that $\mathbf{p} = \Delta(s_0, s_1) \cdot \mathbf{q} + \Delta(s_0, s_2) \cdot \mathbf{r}$. Moreover:*

1. *there is $(s', \mathbf{t}) \in \{(s_1, \mathbf{q}), (s_2, \mathbf{r})\}$ such that \mathbf{t} is a left accumulation point of $f_{s'}$ and $\text{lslope}(f_{s_0}, \mathbf{p}_1) = \text{lslope}(f_{s'}, \mathbf{t}_1)$;*
2. *there is $\eta(s_0, \mathbf{p}) > 0$ such that for all $\varepsilon \in (0, \eta(s_0, \mathbf{p}))$:*
 - *there are $\varepsilon_1 \geq 0, \varepsilon_2 \geq 0$ such that $\text{rslope}(f_{s_0}, \mathbf{p}_1 - \varepsilon) \geq \text{rslope}(f_{s_1}, \mathbf{q}_1 - \varepsilon_1)$, $\text{rslope}(f_{s_0}, \mathbf{p}_1 - \varepsilon) \geq \text{rslope}(f_{s_2}, \mathbf{r}_1 - \varepsilon_2)$, and $\varepsilon = \Delta(s_0, s_1) \cdot \varepsilon_1 + \Delta(s_0, s_2) \cdot \varepsilon_2$;*
 - *if \mathbf{r} is not a left accumulation point in f_{s_2} , or $\text{lslope}(f_{s_0}, \mathbf{p}_1) \neq \text{lslope}(f_{s_2}, \mathbf{r}_1)$, then $f_{s_0}(\mathbf{p}_1 - \varepsilon) = \Delta(s_0, s_1) \cdot f_{s_1}\left(\frac{\mathbf{p}_1 - \varepsilon - \Delta(s_0, s_2) \cdot \mathbf{r}_1}{\Delta(s_0, s_1)}\right) + \Delta(s_0, s_2) \cdot \mathbf{r}_2$;*
 - *symmetrically, if \mathbf{q} is not a left accumulation point in f_{s_1} , or $\text{lslope}(f_{s_0}, \mathbf{p}_1) \neq \text{lslope}(f_{s_1}, \mathbf{q}_1)$, then $f_{s_0}(\mathbf{p}_1 - \varepsilon) = \Delta(s_0, s_1) \cdot \mathbf{q}_2 + \Delta(s_0, s_2) \cdot f_{s_2}\left(\frac{\mathbf{p}_1 - \varepsilon - \Delta(s_0, s_1) \cdot \mathbf{q}_1}{\Delta(s_0, s_2)}\right)$.*

Proof (Sketch). We use the fact that for every extremal point \mathbf{p}' there are unique extremal points \mathbf{q}' and \mathbf{r}' on f_{s_1} and f_{s_2} , respectively, such that $\mathbf{p}' = \Delta(s_0, s_1) \cdot \mathbf{q}' + \Delta(s_0, s_2) \cdot \mathbf{r}'$.

To prove item 1, we show that for all extremal point \mathbf{p}' , $\text{lslope}(s_0, \mathbf{p}') = \min(\text{lslope}(s_1, \mathbf{q}'), \text{lslope}(s_2, \mathbf{r}'))$, which can be surprising at first glance since one could have expected a weighted sum of the left slopes. This fact is illustrated in Fig. 4 (right): $\text{lslope}(s_0, \mathbf{p}') = \text{lslope}(s_1, \mathbf{q}') \leq \text{lslope}(s_2, \mathbf{r}')$. The inequality $\text{lslope}(s_0, \mathbf{p}) \leq \text{lslope}(s_1, \mathbf{q})$ (and similarly $\text{lslope}(s_0, \mathbf{p}) \leq \text{lslope}(s_2, \mathbf{r})$), follows from concavity of f_{s_0} : because for all $\varepsilon > 0$ the inequality $f_{s_0}(\mathbf{p}_1 - \varepsilon) \geq \Delta(s_0, s_1) \cdot f_{s_1}\left(\mathbf{q}_1 - \frac{\varepsilon}{\Delta(s_0, s_1)}\right) + \Delta(s_0, s_2) \cdot f_{s_2}(\mathbf{r}_1)$ holds true, from which we obtain $\lim_{\varepsilon \rightarrow 0^+} \frac{f_{s_0}(\mathbf{p}_1) - f_{s_0}(\mathbf{p}_1 - \varepsilon)}{\varepsilon} \leq \lim_{\varepsilon \rightarrow 0^+} \frac{f_{s_1}(\mathbf{q}_1) - f_{s_1}\left(\mathbf{q}_1 - \frac{\varepsilon}{\Delta(s_0, s_1)}\right)}{\frac{\varepsilon}{\Delta(s_0, s_1)}}$. Showing that the left slope is at least the minimum of the successors' slopes is significantly more demanding and technical, and we give the proof in the appendix.

Proving the second point, is based on the observation that a point on the Pareto curve f_{s_0} is a combination of points of f_{s_1} and f_{s_2} that share a common tangent: in other words they maximize the dot product with a specific vector on their respective curves. From this observation it is possible to link the right slopes of these curves.

The last two points hold because with the assumption, extremal points that converge to \mathbf{p} from the left can be obtained as a combination from a fixed \mathbf{r} and points on f_{s_2} . \square

Now we will prove that there are no left accumulation points on the Pareto curve. To do that, we will try to follow one in the game: if there is a left accumulation point in one state then at least one of its successors also has one, as the

above lemmas show. By using the fact that the left slopes of left accumulation points are preserved we show that the number of reachable combinations (s, \mathbf{p}) , where $s \in S$ and \mathbf{p} is a left accumulation point, is finite. We then look at points slightly to the left of the accumulation points, their distance to the accumulation point and right slopes are also mostly preserved except in stochastic states, where if only one successor has a left accumulation point, the decrease of the right slope accelerate (by Lem. 11.2). By using the fact that in stopping games we can ensure visiting such stochastic states, we will show that for some states the right slope is constant on the left neighbourhood of the left accumulation point, which is a contradiction.

Assume we are given a state s_0 and a left accumulation point \mathbf{p}^0 of f_{s_0} . We construct a transition system T_{s_0, \mathbf{p}^0} where the initial state is (s_0, \mathbf{p}^0) , and the successors of a given configuration (s, \mathbf{p}) are the states (s', \mathbf{p}') such that s' is a successor of s , and \mathbf{p}' is a left accumulation point of s' with the same left slope on $f_{s'}$ as \mathbf{p} on f_s . Lem. 9, 10, and 11, ensure that all the reachable states have at least one successor.

Lemma 12. *For all reachable states (s, \mathbf{p}) and (s', \mathbf{p}') in the transition system T_{s_0, \mathbf{p}^0} , if $s = s'$, then $\mathbf{p} = \mathbf{p}'$.*

Proof. Assume $s = s'$. By construction of T_{s_0, \mathbf{p}^0} , the left slope in s of \mathbf{p} and \mathbf{p}' is the same: $\text{lslope}(s, \mathbf{p}_1) = \text{lslope}(s_0, \mathbf{p}_1^0) = \text{lslope}(s_2, \mathbf{p}_1')$. Assume towards a contradiction that $\mathbf{p} < \mathbf{p}'$; the proof would work the same for $\mathbf{p}' < \mathbf{p}$. Since \mathbf{p}' is a left accumulation point, there is an extremal point in $(\mathbf{p}_1, \mathbf{p}_1')$. Lem. 8.2 tells us that $\text{lslope}(s_1, \mathbf{p}_1) \neq \text{lslope}(s_2, \mathbf{p}_1')$ which is a contradiction. Hence $\mathbf{p} = \mathbf{p}'$. \square

As a corollary of this lemma, the number of states that are reachable in T_{s_0, \mathbf{p}^0} is finite and bounded by $|S|$.

4.2 Inverse betting game

To show a contradiction, we will follow a path with left accumulation points. We want this path to visit stochastic states which have only one successor in T_{s_0, \mathbf{p}^0} . For that, we will prove a property of an intermediary game that we call an inverse betting game.

An *inverse betting game* is a two player game, given by $\langle V_\exists, V_\forall, E, (v_0, c_0), w \rangle$ where V_\exists and V_\forall are the set of vertices controlled by Eve and Adam, respectively, $\langle V_\exists \cup V_\forall, E \rangle$ is a graph whose each vertex has two successors, $(v_0, c_0) \in V \times \mathbb{R}$ is the initial configuration, and $w: E \rightarrow \mathbb{R}$ is a weight function such that for all $v \in V$: $\sum_{v' | (v, v') \in E} w(v, v') = 1$.

A configuration of the game is a pair $(v, c) \in V \times \mathbb{R}$ where v is a vertex and c a credit. The game starts in configuration (v_0, c_0) and is played by two players Eve and Adam. At each step, from a configuration (v, c) controlled by Eve, Adam suggests a valuation $d: E \rightarrow \mathbb{R}$ for the outgoing edges of v such that $\sum_{v' | (v, v') \in E} w(v, v') \cdot d(v, v') = c$. Eve then chooses a successor v' such that $(v, v') \in E$ and the game continues from configuration $(v', d(v, v'))$. From

a configuration (v, c) controlled by Adam, Adam choses a successor v' of v and keeps the same credit, hence the game continues from (v', c) .

Intuitively, Adam has some credit, and at each step he has to distribute it by betting over the possible successors. Then Eve choses the successor and Adam gets a credit equal to its bet divided by the probability of this transition. The game is *inverse* because Eve is trying to maximize the credit of Adam.

Theorem 13. *Let $\langle V_\exists, V_\forall, E, (v_0, c_0), w \rangle$ be an inverse betting game. Let $T \subseteq V_\exists \cup V_\forall$ be a target set and $B \in \mathbb{R}$ a bound. If from every vertex $v \in V$, Eve has a strategy to ensure visiting T then she has one to ensure visiting it with a credit $c \geq 1$ or to exceed the bound, that is, she can force a configuration in $(T \times [c_0, +\infty)) \cup (V \times [B, +\infty))$.*

Our next step is transforming the transition system T_{s_0, \mathbf{p}^0} into such a game. Consider the inverse betting game \mathcal{B} on the structure given by T_{s_0, \mathbf{p}^0} where $V_\exists = S_\circ$ are the states controlled by Eve, $V_\forall = S_\square \cup S_\diamond$ are controlled by Adam, $w((s, \mathbf{p}), (s', \mathbf{p}')) = \Delta(s, s')$ is a weight on edges and the initial configuration is $((s_0, \mathbf{p}^0), \varepsilon_0)$. Let U_{s_0, \mathbf{p}^0} the set of terminal states and of stochastic states that have only one successor in T_{s_0, \mathbf{p}^0} . We show that in the inverse betting game obtained from a stopping game \mathcal{G} , Eve can ensure visiting U_{s_0, \mathbf{p}^0} .

Lemma 14. *If \mathcal{G} is stopping, there is a strategy for Eve in \mathcal{B} such that from every vertex $v \in V$, all outcomes visit U_{s_0, \mathbf{p}^0} .*

Proof. Assume towards a contradiction that this is not the case, then by memoryless determinacy of turn-based reachability games (see e.g. [10]) there is a vertex v and a memoryless deterministic strategy σ_{Adam} of Adam, such that no outcomes of σ_{Adam} from v visit U_{s_0, \mathbf{p}^0} . Let π and σ be the strategies of Player 1 and Player 2 respectively corresponding to σ_{Adam} . Formally, if $h \in \Omega_{\mathcal{G}}^\square$ then $\pi(h) = \sigma_{\text{Adam}}(h)$ and if $h \in \Omega_{\mathcal{G}}^\diamond$ then $\sigma(h) = \sigma_{\text{Adam}}(h)$. We prove that all outcomes λ in \mathcal{G} of π, σ from v are outcomes of σ_{Adam} in \mathcal{B} . This is by induction on the prefixes $\lambda_{\leq i}$ of the outcomes. It is clear when $\lambda_{\leq i}$ ends with states that are controlled by Player 1 and Player 2 by the way we defined π and σ , that $\lambda_{\leq i+1}$ is also compatible with σ_{Adam} in \mathcal{B} . For a finite path $\lambda_{\leq i}$ ending with a stochastic state s in \mathcal{G} , two successors are possible. With the induction hypothesis that $\lambda_{\leq i}$ is compatible with σ_{Adam} , and by the assumption on σ_{Adam} , s does not belong to U_{s_0, \mathbf{p}^0} . Therefore, both successors of s are also in T_{s_0, \mathbf{p}^0} , and $\lambda_{\leq i+1}$ is compatible with σ_{Adam} in \mathcal{B} . This shows that outcomes in \mathcal{G} of (π, σ) are also outcomes of σ_{Adam} in \mathcal{B} . Therefore, π and σ ensure that from v , we visit no state of U_{s_0, \mathbf{p}^0} and thus no terminal state. This contradicts that the game is stopping. \square

Putting Thm. 13 and Lem. 14 together we can conclude the following:

Corollary 15. *If \mathcal{G} is stopping then in \mathcal{B} , for any bound B , Eve has a strategy to ensure visiting U_{s_0, \mathbf{p}^0} with a credit $c \geq 1$ or making c exceed B .*

4.3 Contradicting sequence

We define $\theta(s_0, \mathbf{p}^0) = \min\{\eta(s, \mathbf{p}) \mid (s, \mathbf{p}) \text{ reachable in } T_{s_0, \mathbf{p}^0}\}$, and consider a sequence of points that are $\theta(s_0, \mathbf{p}^0)$ close to \mathbf{p}^0 and with a right slope that is decreasing at least as fast as that of their predecessors.

Lemma 16. *For stopping games, given $s_0 \in S$, $\mathbf{p}^0 \in \mathbb{R}^2$, and $\varepsilon_0 > 0$, such that $\varepsilon_0 < \theta(s_0, \mathbf{p}^0)$, there is a finite sequence $\pi(s_0, \mathbf{p}^0, \varepsilon_0) = (s_i, \mathbf{p}^i, \varepsilon_i)_{i \leq j}$ such that:*

- $(s_i, \mathbf{p}^i)_{i \leq j}$ is a path in T_{s_0, \mathbf{p}^0} ;
- for all $i \leq j$, $\text{rslope}(f_{s_i}, \mathbf{p}_1^i - \varepsilon_i) \geq \text{rslope}(f_{s_{i+1}}, \mathbf{p}_1^{i+1} - \varepsilon_{i+1})$.
- either $\varepsilon_j \geq \theta(s_0, \mathbf{p}^0)$ or $s_j \in U_{s_0, \mathbf{p}^0}$ and $\varepsilon_j \geq \varepsilon_0$.

The idea of the proof is that in \mathcal{B} , thanks to Lem. 9, 10, and 11, Adam can always choose a successor such that $\text{rslope}(f_{s_i}, \mathbf{p}_1^i - \varepsilon_i) \geq \text{rslope}(f_{s_{i+1}}, \mathbf{p}_1^{i+1} - \varepsilon_{i+1})$. Then thanks to Cor. 15, there is a strategy for Eve to reach $(U_{s_0, \mathbf{p}^0} \times [c_0, +\infty)) \cup (V \times [B, +\infty))$. By combining the two strategies, we obtain an outcome that satisfies the desired properties.

We use the path obtained from this lemma to show that no matter how small ε_0 we choose, ε_i can grow to reach $\theta(s_0, \mathbf{p}^0)$.

Lemma 17. *For all states s with a left accumulation point \mathbf{p} and for all $0 < \varepsilon < \theta(s, \mathbf{p})$, there is some (s', \mathbf{p}') reachable in $T_{s, \mathbf{p}}$ such that $\text{rslope}(f_{s'}, \mathbf{p}_1' - \theta(s, \mathbf{p})) \leq \text{rslope}(f_s, \mathbf{p}_1 - \varepsilon)$.*

Thanks to this lemma, we can now prove Thm. 6. Assume towards a contradiction that there is a left accumulation point \mathbf{p} in the state s . Let $m = \min\{\text{lslope}(f_{s'}, \mathbf{p}_1' - \theta(s, \mathbf{p})) \mid (s', \mathbf{p}') \text{ reachable in } T_{s, \mathbf{p}}\}$ and (s', \mathbf{p}') the configuration of $T_{s, \mathbf{p}}$ for which this minimum is reached (it is reached because the number of reachable configurations is finite: this is a corollary of Lem. 12). Because of Lem. 17, $\text{rslope}(f_s, \mathbf{p}_1 - \varepsilon)$ is greater than m . By Lem. 8.3, when ε goes towards 0, $\text{rslope}(f_s, \mathbf{p}_1 - \varepsilon)$ converges to $\text{lslope}(f_s, \mathbf{p}_1)$. This means that $\text{lslope}(f_s, \mathbf{p}_1) \geq m$. Moreover, by construction of $T_{s, \mathbf{p}}$, we also have that $\text{lslope}(f_{s'}, \mathbf{p}_1') = \text{lslope}(f_s, \mathbf{p}_1)$, so $\text{lslope}(f_{s'}, \mathbf{p}_1') \geq m$. Because the slopes are decreasing (Lem. 8.1), $m = \text{rslope}(f_{s'}, \mathbf{p}_1' - \theta(s, \mathbf{p})) \geq \text{lslope}(f_{s'}, \mathbf{p}_1') \geq m$. Hence, the left and right slopes of $f_{s'}$ are constant on $[\mathbf{p}_1' - \theta(s, \mathbf{p}), \mathbf{p}_1']$, and Lem. 8.2 implies that there are no extremal point in $(\mathbf{p}_1' - \theta(s, \mathbf{p}), \mathbf{p}_1')$. This contradicts the fact that \mathbf{p}' is a left accumulation point: there should be an extremal point in any neighbourhood on the left of \mathbf{p}' . Hence, f_s contains no accumulation point.

Remark 18. One might attempt to extend the proof of Thm. 6 to three or more objectives, but this does not seem to be easily doable. Although it is possible to use directional derivative (or pick a subgradient) instead of using left and right slope in such setting, an analogue of Lem. 8.2 cannot be proved because in multiple dimensions, two accumulation points can share the same directional derivative, for a fixed direction. It is also not easily possible to avoid this problem by following several directional derivatives instead of just one. This is because the slope in one direction may be inherited from one successor while the slope

in another direction comes from another successor. We give more details and example of convex sets that would contradict generalisations of Lem. 8.2 and Lem. 10 in Appendix E.

5 Conclusions

We have studied stochastic games under multiple objectives, and have provided decidability results for determined games and for stopping games with two objectives. Our results for non-determined games provide an important milestone towards obtaining decidability for the general case, which is a major task which will require further novel insights into the problem. Another research direction concerns establishing an upper bound on the number of extremal points of a Pareto curve; such result would allow us to give upper complexity bounds for the problem.

Acknowledgements. The authors would like to thank Aistis Šimaitis and Clemens Wiltsche for their useful discussions on the topic.

References

1. N. Basset, M. Kwiatkowska, U. Topcu, and C. Wiltsche. Strategy synthesis for stochastic games with multiple long-run objectives. In *TACAS*. Springer, 2015.
2. R. Brenguier and J. Raskin. Pareto curves of multidimensional mean-payoff games. In *CAV*, 2015.
3. K. Chatterjee and L. Doyen. Perfect-information stochastic games with generalized mean-payoff objectives. In *LICS*, 2016. To appear.
4. K. Chatterjee and T. A. Henzinger. A survey of stochastic ω -regular games. *J. Comput. Syst. Sci.*, 78(2), 2012.
5. K. Chatterjee, M. Randour, and J. Raskin. Strategy synthesis for multi-dimensional quantitative objectives. *Acta Inf.*, 51(3-4), 2014.
6. T. Chen, V. Forejt, M. Kwiatkowska, A. Simaitis, A. Trivedi, and M. Ummels. Playing stochastic games precisely. In *CONCUR*, 2012.
7. T. Chen, V. Forejt, M. Kwiatkowska, A. Simaitis, and C. Wiltsche. On stochastic games with multiple objectives. In *MFCS*. Springer, 2013.
8. A. Condon. The complexity of stochastic games. *Inf. Comput.*, 96(2), 1992.
9. K. Etessami, M. Kwiatkowska, M. Vardi, and M. Yannakakis. Multi-objective model checking of Markov decision processes. *LMCS*, 4(4), 2008.
10. E. Grädel, W. Thomas, and T. Wilke. *Automata, Logics, and Infinite Games: A Guide to Current Research*, volume 2500. Springer, 2003.
11. P. Hunter and J. Raskin. Quantitative games with interval objectives. In *FSTTCS*, 2014.
12. K. G. Larsen, M. Mikucionis, M. Muñoz, J. Srba, and J. H. Taankvist. Online and compositional learning of controllers with application to floor heating. In *TACAS*, 2016.
13. J. F. Nash, Jr. Equilibrium points in n -person games. *Proc. National Academy of Sciences of the USA*, 36(1), Jan. 1950.
14. A. Tarski. *A Decision Method for Elementary Algebra and Geometry*. Univ. of California Press, Berkeley, 1951.

15. Y. Velner. Robust multidimensional mean-payoff games are undecidable. In *FoSSaCS*. Springer, 2015.
16. Y. Velner, K. Chatterjee, L. Doyen, T. A. Henzinger, A. M. Rabinovich, and J. Raskin. The complexity of multi-mean-payoff and multi-energy games. *Inf. Comput.*, 241, 2015.

A Proofs for determined games

A.1 Proof of Lem. 5

We write Σ_{MD} for the set of memoryless deterministic strategies of Player 2.

$$\begin{aligned}
& \forall \varepsilon > 0. \exists \pi \in \Pi. \forall \sigma \in \Sigma. \bigwedge_i \mathbb{E}^{\pi, \sigma}(\mathbf{z}_i) \geq \mathbf{z}_i - \varepsilon \\
\Leftrightarrow & \quad \forall \varepsilon > 0. \neg \exists \sigma \in \Sigma. \forall \pi \in \Pi. \bigvee_i \mathbb{E}^{\pi, \sigma}(\mathbf{z}_i) < \mathbf{z}_i - \varepsilon \quad (\text{determinacy}) \\
\Leftrightarrow & \quad \forall \varepsilon > 0. \neg \exists \sigma \in \Sigma_{\text{MD}}. \forall \pi \in \Pi. \bigvee_i \mathbb{E}^{\pi, \sigma}(\mathbf{z}_i) < \mathbf{z}_i - \varepsilon \quad (\text{Thm. 4}) \\
\Leftrightarrow & \quad \forall \varepsilon > 0. \forall \sigma \in \Sigma_{\text{MD}}. \exists \pi \in \Pi. \bigwedge_i \mathbb{E}^{\pi, \sigma}(\mathbf{z}_i) \geq \mathbf{z}_i - \varepsilon \quad (\text{logical equivalences}) \\
\Leftrightarrow & \quad \forall \sigma \in \Sigma_{\text{MD}}. \forall \varepsilon > 0. \exists \pi \in \Pi. \bigwedge_i \mathbb{E}^{\pi, \sigma}(\mathbf{z}_i) \geq \mathbf{z}_i - \varepsilon \quad (\text{logical equivalences}) \\
\Leftrightarrow & \quad \forall \sigma \in \Sigma_{\text{MD}}. \exists \pi \in \Pi. \bigwedge_i \mathbb{E}^{\pi, \sigma}(\mathbf{z}_i) \geq \mathbf{z}_i \quad (\text{property of MDPs [9]})
\end{aligned}$$

B Proof for non-determined case

Lemma 19. *Let s be a state of \mathcal{G} .*

1. *The domain of definition of f_s is an interval of \mathbb{R} .*
2. *f_s is decreasing and concave.*
3. *If f_s is defined on (a, b) then f_s is continuous on (a, b) .*

Proof. 1. Assume that f_s is defined for some x, x' . For all $x'' \in (x, x')$, by convexity $(x'', \frac{x''-x'}{x-x'} \cdot f_s(x) + \frac{x''-x}{x'-x} \cdot f_s(x'))$ can be achieved. We define $y = \sup\{y \mid (x'', y) \text{ achievable}\}$, it is defined because of the previous remark, and by the properties of stopping games (x'', y) is achievable and we now show that $(x'', y) \in f_s$.

Assume there is an achievable \mathbf{p} with $\mathbf{p}_1 > x''$. Then by convexity of the set of achievable points [7], any point in $[(x, f_s(x)), \mathbf{p}]$ is achievable. Because $(x, f_s(x))$ lies on the Pareto curve, we have $f_s(x) > \mathbf{p}_2$ (otherwise $(x, f_s(x))$ is not maximal point satisfying the defining properties of the Pareto curve), therefore $(x'', \frac{x''-\mathbf{p}_1}{x-\mathbf{p}_1} \cdot f_s(x) + \frac{x''-x}{\mathbf{p}_1-x} \cdot f_s(\mathbf{p}_1))$ is achievable which contradicts the definition of y . Therefore, f_s is defined in x'' and equals \mathbf{p}_2 .

2. The fact that if $x' \geq x$, then $f_s(x) \geq f_s(x')$ comes by the definition of Pareto curves and the fact that $(x, f_s(x))$ lies on a Pareto curve. We now prove that f_s is concave. Let $x < x'$ in the domain of f_s and $t \in [0, 1]$. By convexity of the set \mathcal{A}_s [7], $t \cdot (x, f_s(x)) + (1-t) \cdot (x', f_s(x'))$ can be ensured. Hence, there is a point \mathbf{p} that is strictly greater than $t \cdot (x, f_s(x)) + (1-t) \cdot (x', f_s(x'))$ that belongs to f_s . Since f_s is defined in $t \cdot x + (1-t) \cdot x'$ and it is decreasing this means that $f_s(t \cdot x + (1-t) \cdot x') \geq t \cdot f_s(x) + (1-t) \cdot f_s(x')$. This shows concavity.

3. Since f_s is concave, it is the negation of a convex function. A convex function defined on an open interval is continuous on this interval. Therefore, if f_s is defined on (a, b) it is continuous on (a, b) . \square

B.1 Proof of Lem. 8

We prove a bit more than what Lem. 8 gives in the main part of the paper:

- Lemma 20.** 1. Let f be a function (not necessarily concave) whose left slope is well defined on the interval $[a, b]$, then there exists $x \in [a, b]$ such that $\text{rslope}(f, x) \geq \frac{f(a)-f(b)}{a-b}$. Similarly, there exists $x' \in [a, b]$ such that $\text{lslope}(f, x') \leq \frac{f(a)-f(b)}{a-b}$.
2. If f is concave, then $x \mapsto \text{lslope}(f, x)$ is defined on the same interval as f except the left most point.
3. If f is concave and $x < x'$ are two reals for which $\text{lslope}(f)$ is defined, then $\text{lslope}(f, x) \geq \text{rslope}(f, x) \geq \text{lslope}(f, x')$.
4. If f is concave and (x, x') contains an extremal point of f , then $\text{lslope}(f, x) \neq \text{lslope}(f, x')$.
5. If f is concave and $\text{lslope}(f, x) \neq \text{lslope}(f, x')$ then $[x, x']$ contains an extremal point.
6. If f is a concave function whose left slope is defined in x , then $\lim_{x' \rightarrow x^-} \text{lslope}(f, x') = \lim_{x' \rightarrow x^-} \text{rslope}(f, x') = \text{lslope}(f, x)$.

Proof. 1. First we show that if $\text{rslope}(f, x) < d$ for all $x \in [a, b]$ then for all $x \in (a, b]$, $x < f(a) + (x - a) \cdot d$. To prove this, let c be the right-most point of $[a, b]$ which minimizes $f(c) - f(a) - (c - a) \cdot d$. Assume towards a contradiction that $c \neq b$. Since $\text{rslope}(f, c) < d$, there exists $x \in (c, b)$ such that $\frac{f(x)-f(c)}{x-c} < d$. Since $(x - c)$ is positive this implies that:

$$\begin{aligned} f(x) &< f(c) + (x - c) \cdot d \\ f(x) - f(a) - (x - a) \cdot d &< f(c) - f(a) - (x - a - x + c) \cdot d \\ f(x) - f(a) - (x - a) \cdot d &< f(c) - f(a) - (c - a) \cdot d \end{aligned}$$

Which contradicts the fact that c is the right-most point of $[a, b]$ maximizing $f(c) - f(a) - (c - a) \cdot d$. Hence, b is the only point that maximizes this quantity, and it equals 0 in b . We therefore have that for all $x \in [a, b]$, $f(x) < f(a) + (x - a) \cdot d$.

In particular, assuming towards a contradiction that for all $x \in [a, b]$, $\text{rslope}(f, x) < \frac{f(a)-f(b)}{a-b}$, we would have $f(a) < f(b) + (a - b) \cdot \frac{f(a)-f(b)}{a-b} = f(a)$. Hence, there exists some $x \in [a, b]$ such that $\text{rslope}(f, x) \geq \frac{f(a)-f(b)}{a-b}$.

Consider now the function $-f$. As a consequence of what we just proved, there is $x \in [a, b]$ such that $\text{lslope}(-f, x) \geq \text{rslope}(-f, x) \geq -\frac{f(a)-f(b)}{a-b}$. Since $\text{lslope}(-f, x) = -\text{lslope}(f, x)$, this implies that $\text{lslope}(f, x) \leq \frac{f(a)-f(b)}{a-b}$.

2. Assume that f is defined on $[a, b]$ and let $x \in (a, b]$. Define the function g by $g(x') = \frac{f(x')-f(x)}{x'-x}$, it is defined for x' smaller than x and close enough to x .

Moreover by concavity of f , g is decreasing, therefore its left limit in x is well defined, and equals the left slope of f in x .

3. By concavity, for $x < x'$ and $0 < \varepsilon < \frac{x'-x}{2}$:

$$\frac{f(x) - f(x - \varepsilon)}{\varepsilon} \geq \frac{f(x) - f(x + \varepsilon)}{\varepsilon} \geq \frac{f(x') - f(x' - \varepsilon)}{\varepsilon}$$

Hence, it is the same for the limit when ε moves towards 0 and $\text{lslope}(f, x) \geq \text{lslope}(f, x')$.

4. Assume towards a contradiction that $\text{lslope}(f, x) = \text{lslope}(f, x')$, and \mathbf{p} is an extremal point with $x < \mathbf{p}_1 < x'$. By item 3 this means the slope is constant on $[x, x']$ and as a consequence of item 1, it is equal to $\frac{f(x') - f(x)}{x' - x}$. Since it is an extremal point, $f(\mathbf{p}_1) > f(x) + (\mathbf{p}_1 - x) \cdot \frac{f(x') - f(x)}{x' - x}$. By item 1, there is $x'' \in [x, \mathbf{p}_1]$ such that:

$$\begin{aligned} \text{lslope}(f, x'') &\geq \frac{f(\mathbf{p}_1) - f(x)}{\mathbf{p}_1 - x} \\ &> \frac{f(x) + (\mathbf{p}_1 - x) \cdot \frac{f(x') - f(x)}{x' - x} - f(x)}{\mathbf{p}_1 - x} \\ &> \frac{(\mathbf{p}_1 - x) \cdot \frac{f(x') - f(x)}{x' - x}}{\mathbf{p}_1 - x} \\ &> \frac{f(x') - f(x)}{x' - x} \end{aligned}$$

Which is in contradiction with the fact that on $[x, x']$ the slope is constant and equal to $\frac{f(x') - f(x)}{x' - x}$.

5. Assume towards a contradiction that $[x, x']$ contains no extremal point, then for all $x'' \in [x, x']$, $(x'', f(x''))$ is a convex combination of $(x, f(x))$ and $(x', f(x'))$. Hence, if $x'' \in (x, x']$, $\text{lslope}(f, x'') = \frac{f(x) - f(x'')}{x - x''}$. This is in particular the case for x' . Since $\text{lslope}(f, x) \neq \text{lslope}(f, x')$, x has a left slope different from all the points in $(x, x']$; we will prove that $(x, f(x))$ is an extremal point. If x is a convex combination of two points of the curve given by f , then as all the points on its right are below the line segment $[(x, f(x)), (x', f(x'))]$, one point $\mathbf{p} \in f$ with $\mathbf{p}_1 < x$ must be above the corresponding line. By concavity, this is also the case for all the $x'' \in (\mathbf{p}_1, x)$. This means that $\frac{f(x'') - f(x)}{x'' - x} \leq \frac{f(x') - f(x)}{x' - x}$. The limit when x'' goes towards x is also smaller than $\frac{f(x') - f(x)}{x' - x}$. It cannot be greater because by item 3 the slope decreases. This contradicts that $\text{lslope}(f, x) \neq \text{lslope}(f, x')$. This contradiction proves that $[x, x']$ contains an extremal point.

6. The fact that $\lim_{x' \rightarrow x^-} \text{lslope}(f, x') \geq \lim_{x' \rightarrow x^-} \text{rslope}(f, x') \geq \text{lslope}(f, x)$ comes from the fact that the slope is decreasing (point 3). We now show $\text{lslope}(f, x) \leq \lim_{x' \rightarrow x^-} \text{lslope}(f, x')$ which implies the result. Assume towards a contradiction that $\lim_{x' \rightarrow x^+} \text{lslope}(f, x') > \text{lslope}(f, x)$. Then there is $\varepsilon > 0$ such that $\frac{f(x) - f(x - \varepsilon)}{\varepsilon} < \lim_{x' \rightarrow x^-} \text{lslope}(f, x')$. By point 1, there is $x'' \in [x - \varepsilon, x]$ such that $\text{lslope}(f, x'') \leq \frac{f(x) - f(x - \varepsilon)}{\varepsilon} < \lim_{x' \rightarrow x^-} \text{lslope}(f, x')$. Since the slope is decreasing, $\lim_{x' \rightarrow x^-} \text{lslope}(f, x') \geq \text{lslope}(f, x'')$ which gives a contradiction. \square

B.2 Evolution of the slope in Player 1 states (Proof of Lem. 9)

Lemma 21. *Let s_0 be a Player 1 state with two successors s_1 and s_2 . For all x where the slopes of f_{s_1} and f_{s_2} are defined, f_{s_0} and its slope are defined and:*

$$\begin{aligned} \min(\text{lslope}(f_{s_1}, x), \text{lslope}(f_{s_2}, x)) &\leq \text{lslope}(f_{s_0}, x) \leq \max(\text{lslope}(f_{s_1}, x), \text{lslope}(f_{s_2}, x)) \\ \min(\text{rslope}(f_{s_1}, x), \text{rslope}(f_{s_2}, x)) &\leq \text{rslope}(f_{s_0}, x) \leq \max(\text{rslope}(f_{s_1}, x), \text{rslope}(f_{s_2}, x)) \end{aligned}$$

Proof. To lighten the notation we will write f_i instead of f_{s_i} in the following.

Assume towards a contradiction that there is x where f_1 and f_2 are defined and where f_0 is not defined. Let y be the maximum such that $\mathbf{p} = (x, y)$ can be ensured from s_0 . Since f_0 is not defined in x , there exists $(x', y') \in \mathcal{A}_{s_0}$ such that $x < x'$ and $y \leq y'$. By downward closure of the set of achievable points, (x, y') can be ensured from s_0 , thus by definition of y , $y' \leq y$ and therefore $y = y'$. Then there is $(x', y) \in \mathcal{A}_{s_0}$ with $x' > x$. By the characterisation of the set of values that can be ensured, we have $(x', y) \in \text{conv}(\mathcal{A}_{s_1} \cup \mathcal{A}_{s_2})$, so there are $\lambda_1, \lambda_2 \in [0, 1]$ and $\mathbf{q}, \mathbf{r} \in \mathcal{A}_{s_1} \cup \mathcal{A}_{s_2}$ such that $\lambda_1 + \lambda_2 = 1$ and $(x', y) = \lambda_1 \cdot \mathbf{q} + \lambda_2 \cdot \mathbf{r}$. If $\mathbf{q}_1 > y$ then there is some \mathbf{p}' which can be ensured with $\mathbf{p}'_1 > x$ and $\mathbf{p}'_2 > y'$ which by downward closedness contradicts the fact that y is the maximum such that (x, y) can be ensured. The case $\mathbf{r}_2 > y$ is similar. In the remaining cases $y = \mathbf{q}_2 = \mathbf{r}_2$. Assume w.l.o.g. $\mathbf{q}_1 \geq x$, and therefore $\mathbf{q} \geq (x, y)$. Since f_1 is defined in x , $f_1(x) > \mathbf{q}_2$. By characterisation of the set of values that can be ensured $y \geq f_1(x) > \mathbf{q}_2$ which is a contradiction.

This shows that if f_1 and f_2 are defined then f_0 also is. Since $\text{lslope}(f_0, x)$ is defined for all x where f_0 except the left-most point, it is defined where f_1 and f_2 are defined except the left-most point of one of the two. In this left-most point one of the two slopes is not defined. This shows that if both left slopes are defined then $\text{lslope}(f_0, x)$ also is. The proof proceeds similarly for the right slopes.

Let $m = \min(\text{lslope}(f_1, x), \text{lslope}(f_2, x))$ and $\mathbf{n} = (-m, 1)$. Assume towards a contradiction that $\text{lslope}(f_0, x) < m$. Then by definition of lslope there exists $x' < x$ such that $\frac{f_0(x) - f_0(x')}{x - x'} < m$. Since $x - x' > 0$, we have $f_0(x') > f_0(x) + (x' - x) \cdot m$. Therefore, $(x', f_0(x')) \cdot \mathbf{n} = f_0(x') - mx' > f_0(x) - mx = (x, f_0(x)) \cdot \mathbf{n}$. By the characterization of the Pareto curve, there are $\mathbf{q}', \mathbf{r}' \in \mathcal{A}_{s_1} \cup \mathcal{A}_{s_2}$, such that $(x', f_0(x'))$ is a convex combination of \mathbf{q}' and \mathbf{r}' . Because their convex combination has greater dot product with \mathbf{n} than $(x, f_0(x))$, it is also the case of one of the points \mathbf{q}' or \mathbf{r}' .

Without loss of generality, we assume it is \mathbf{q}' . So we have $f_1(\mathbf{q}'_1) > f_0(x) + (\mathbf{q}'_1 - x) \cdot m$.

- If $\mathbf{q}'_1 < x$, then $\frac{f_0(x) - f_1(\mathbf{q}'_1)}{x - \mathbf{q}'_1} < m \leq \text{lslope}(f_1, x)$. By the characterization of the Pareto curve $f_1(x) \leq f_0(x)$, therefore $\frac{f_1(x) - f_1(\mathbf{q}'_1)}{x - \mathbf{q}'_1} < \text{lslope}(f_1, x)$. By Lem. 20.1, there is $x'' \in [\mathbf{q}'_1, x]$ such that $\text{lslope}(f_1, x'') \leq \frac{f_1(x) - f_1(\mathbf{q}'_1)}{x - \mathbf{q}'_1} < \text{lslope}(f_1, x)$. By Lem. 20.3, $\text{lslope}(f_1, x'') \geq \text{lslope}(f_1, x)$ which is a contradiction.

- If $q'_1 \geq x$, then $x \in [x', q'_1]$ and by the characterization of the Pareto curve $f_0(q'_1) \geq f_1(q'_1)$, so $f_0(x'_1) > f_0(x) + (q'_1 - x) \cdot m$. This implies $(q'_1, f_0(q'_1)) \cdot \mathbf{n} > (x, f_0(x)) \cdot \mathbf{n}$. By concavity of f_0 :

$$f_0(x) \geq f_0(x') + (x - x') \cdot \frac{f_0(q'_1) - f_0(x')}{q'_1 - x'}$$

$$(x, f_0(x)) \cdot \mathbf{n} \geq \left(1 - \frac{x - x'}{q'_1 - x'}\right) (x', f_0(x')) \cdot \mathbf{n} + \frac{x - x'}{q'_1 - x'} (q'_1, f_0(q'_1)) \cdot \mathbf{n}$$

We obtain $(x, f_0(x)) \cdot \mathbf{n} > \left(1 - \frac{x - x'}{q'_1 - x'}\right) (x, f_0(x)) \cdot \mathbf{n} + \frac{x - x'}{q'_1 - x'} (x, f_0(x)) \cdot \mathbf{n} = (x, f_0(x)) \cdot \mathbf{n}$, which is a contradiction.

We now turn to the case of \max . By the characterization of the Pareto curve, there are $\mathbf{q}, \mathbf{r} \in \mathcal{A}_{s_1} \cup \mathcal{A}_{s_2}$ such that $(x, f_0(x))$ is a convex combination of \mathbf{q} and \mathbf{r} . Let $f', f'' \in \{f_1, f_2\}$, be such that $q_2 \leq f'(q_1)$ and $r_2 \leq f''(r_1)$. Let also $\lambda_1 \in [0, 1]$ be such that $(x, f_0(x)) = \lambda_1 \cdot \mathbf{q} + (1 - \lambda_1) \cdot \mathbf{r}$. We first prove $\text{lslope}(f_0, x) \leq \text{lslope}(f', q_1)$ (and obtain similarly $\text{lslope}(f_0, x) \leq \text{lslope}(f'', r_1)$). By the characterisation of the Pareto curve, for all $\varepsilon \neq 0$, $f_0(x - \varepsilon) \geq \lambda_1 \cdot f'(q_1 - \frac{\varepsilon}{\lambda_1}) + (1 - \lambda_1) \cdot f''(r_1)$. So for all $\varepsilon > 0$, $\frac{f_0(x) - f_0(x - \varepsilon)}{\varepsilon} \leq \lambda_1 \cdot \frac{f'(q_1) - f'(q_1 - \frac{\varepsilon}{\lambda_1})}{\frac{\varepsilon}{\lambda_1}}$. The limit when ε goes towards 0 is smaller than that of $\frac{f'(q_1) - f'(q_1 - \varepsilon)}{\varepsilon}$. Hence, $\text{lslope}(f_0, x) \leq \text{lslope}(f', q_1)$.

Now, since their convex combination contains x , one of q_1 and r_1 is greater than x . Assume without loss of generality that it is q_1 . Then $q_1 \geq x$ and by Lem. 3, $\text{lslope}(f', x) \geq \text{lslope}(f', q_1) \geq \text{lslope}(f_0, x)$. This shows $\text{lslope}(f_0, x) \leq \max(\text{lslope}(f_1, x), \text{lslope}(f_2, x))$.

We now turn to the case of rslope . This is in fact the same proof if we consider the function $g_0: x \mapsto f_0(1 - x)$, which is also concave and for which $\text{lslope}(g_0, x) = -\text{rslope}(f_0, 1 - x)$: at no point did we use the fact that the slope was negative. So what we proved for f_0 is also valid for g_0 , which means $\min(\text{rslope}(f_{s_1}, x), \text{rslope}(f_{s_2}, x)) \leq \text{rslope}(f_{s_0}, x) \leq \max(\text{rslope}(f_{s_1}, x), \text{rslope}(f_{s_2}, x))$. \square

We are now ready to prove Lem. 9.

Lemma 9. *Let s_0 be a Player 1 state with two successors s_1 and s_2 , and let \mathbf{p} be a left accumulation point of f_{s_0} . Then there is $\eta(s_0, \mathbf{p}) > 0$ such that for all $\varepsilon \in (0, \eta(s_0, \mathbf{p}))$, there is $s' \in \{s_1, s_2\}$ such that: 1. \mathbf{p} is a left accumulation point in $f_{s'}$; 2. $\text{lslope}(s_0, \mathbf{p}_1) = \text{lslope}(s', \mathbf{p}_1)$; 3. $f_{s_0}(\mathbf{p}_1 - \varepsilon) \geq f_{s'}(\mathbf{p}_1 - \varepsilon)$ and $\text{rslope}(f_{s_0}, \mathbf{p}_1 - \varepsilon) \geq \text{rslope}(f_{s'}, \mathbf{p}_1 - \varepsilon)$.*

Proof. By the characterisation of the Pareto curve, the extremal points of the Pareto curve f_{s_0} are included in the extremal points of f_{s_1} and f_{s_2} . Let $(\mathbf{p}^i)_{i \in \mathbb{N}}$ be a sequence of extremal points of the Pareto curve in s_0 which converges to \mathbf{p} . We assume that the first coordinate of the sequence is increasing (note that we can always extract a sub-sequence which satisfies that). Each \mathbf{p}^i is either

an extremal point of s_1 or of s_2 , therefore the Pareto curve of one of the two contains infinitely many points \mathbf{p}^i . We first show that for any s' for which this holds, the first two points are true for s' . We will then show that the third point is satisfied by one such s' .

1. Since there is a subsequence of \mathbf{p}^i which are extremal points of s' and which converges to \mathbf{p} from the left, \mathbf{p} is a left accumulation point of s' .

2. Moreover since $\lim_{x' \rightarrow \mathbf{p}_1^-} \frac{f_{s_0}(x') - f_{s_0}(\mathbf{p}_1)}{x' - \mathbf{p}_1}$ is well defined and $(\mathbf{p}_1^i)_{i \in \mathbb{N}}$ converges to \mathbf{p}_1 , we have that:

$$\begin{aligned} \text{lslope}(s_0, \mathbf{p}_1) &= \lim_{x' \rightarrow \mathbf{p}_1^-} \frac{f_{s_0}(x') - f_{s_0}(\mathbf{p}_1)}{x' - \mathbf{p}_1} \\ &= \lim_{i \rightarrow \infty} \frac{f_{s_0}(\mathbf{p}_1^i) - f_{s_0}(\mathbf{p}_1)}{\mathbf{p}_1^i - \mathbf{p}_1} \\ &= \lim_{i \rightarrow \infty} \frac{f_{s'}(\mathbf{p}_1^i) - f_{s'}(\mathbf{p}_1)}{\mathbf{p}_1^i - \mathbf{p}_1} \\ &= \text{lslope}(s', \mathbf{p}_1) \end{aligned}$$

3. Let $x' < \mathbf{p}_1$. If \mathbf{p} is not a left accumulation point for s_1 , let $x_1 = \sup\{x_1 < \mathbf{p}_1 \mid (x_1, f_{s_1}(x_1)) \text{ extremal in } f_{s_1}\}$ and let \mathbf{q} be an extremal point of s_0 with $\mathbf{q}_1 \in (x_1, \mathbf{p}_1)$. Since extremal points of s_0 are included in those of s_1 and s_2 , the extremal points on $[\mathbf{q}_1, x]$ and x are the same. Since moreover $(\mathbf{q}_1, f_0(\mathbf{q}_1))$ and $(x, f_0(x))$ are extremal points of both f_0 and f_2 , the curves f_0 and f_2 coincide on $[\mathbf{q}_1, x]$. So the property we want is satisfied by $s' = s_2$ and $\eta(s_0, \mathbf{p}) = x - \mathbf{q}_1$.

Similarly, if \mathbf{p} is not a left accumulation point for s_2 , then $s' = s_1$ and some $\eta(s_0, \mathbf{p})$ witnesses the property.

If \mathbf{p} is a left accumulation point for s_1 but $\text{lslope}(f_{s_1}, \mathbf{p}) \neq \text{lslope}(f_0, \mathbf{p})$, then by what was proven in point 2, we conclude that there is no infinite sequence of extremal points of s_1 that are also extremal points of s_0 and converge to \mathbf{p} . There is a left neighbourhood of x where f_1 is below f_2 (otherwise some extremal point would be above and also be an extremal point of f_0). Hence, the curves f_0 and f_{s_2} coincide on some neighbourhood of x . So the property we want is satisfied by $s' = s_2$ and some $\eta(s_0, \mathbf{p})$.

Similarly, if $\text{lslope}(f_{s_1}, \mathbf{p}) \neq \text{lslope}(f_s, \mathbf{p})$, then $s' = s_1$ and some $\eta(s_0, \mathbf{p})$ witnesses the property.

In the other cases, both s_1 and s_2 satisfy points 1 and 2. By Lem. 21, there is $s' \in \{s_1, s_2\}$, such that $\text{rslope}(f_{s'}, x') \leq \text{rslope}(f_{s_0}, x')$. So the property is satisfied for any $\eta(s_0, \mathbf{p})$ such that $x - \eta(s_0, \mathbf{p})$ is still in the domain of f_{s_0} . \square

B.3 Evolution of the slope in Player 2 states (proof of Lem. 10)

Lemma 23. *Let s_0 be a Player 2 state with two successors s_1 and s_2 . For all x where f_{s_1} is defined and f_{s_2} is defined, f_{s_0} is defined in $x + \mathbf{q}_1(s_0)$ and: 1. if $f_{s_1}(x) < f_{s_2}(x)$ then $\text{lslope}(f_{s_0}, x + \mathbf{q}_1(s_0)) = \text{lslope}(f_{s_1}, x)$ and $\text{rslope}(f_{s_0}, x + \mathbf{q}_1(s_0)) = \text{rslope}(f_{s_1}, x)$; 2. if $f_{s_1}(x) = f_{s_2}(x)$ then $\text{lslope}(f_{s_0}, x + \mathbf{q}_1(s_0)) = \max\{\text{lslope}(f_{s_1}, x), \text{lslope}(f_{s_2}, x)\}$ and $\text{rslope}(f_{s_0}, x) = \min\{\text{rslope}(f_{s_1}, x), \text{rslope}(f_{s_2}, x)\}$.*

Proof. 1. Assume $f_{s_1}(x) < f_{s_2}(x)$. By continuity of the curves (Lem. 19.3), $f_{s_1}(x') < f_{s_2}(x')$ holds for all points x' of some neighbourhood $[x - \eta, x + \eta]$. By the characterization of the Pareto curve in **Player 2** states (see Sec. 2.1), we have $f_{s_0}(x' + \boldsymbol{\varrho}_1(s_0)) = \boldsymbol{\varrho}_2(s_0) + \min\{f_{s_1}(x'), f_{s_2}(x')\}$. Hence, for all $x' \in [x - \eta, x + \eta]$, $f_{s_0}(x' + \boldsymbol{\varrho}_1(s_0)) = \boldsymbol{\varrho}_2(s_0) + f_{s_1}(x')$. So, $\text{lslope}(f_{s_0}, x + \boldsymbol{\varrho}_1(s_0)) = \lim_{x' \rightarrow x^-} \frac{\boldsymbol{\varrho}_2(s_0) + f_{s_0}(x') - f_{s_0}(x) - \boldsymbol{\varrho}_2(s_0)}{x' - x} = \lim_{x' \rightarrow x^-} \frac{f_{s_1}(x') - f_{s_1}(x)}{x' - x} = \text{lslope}(f_{s_1}, x)$. Similarly, $\text{rslope}(f_{s_0}, x + \boldsymbol{\varrho}_1(s_0)) = \lim_{x' \rightarrow x^+} \frac{f_{s_0}(x') - f_{s_0}(x)}{x' - x} = \lim_{x' \rightarrow x^+} \frac{f_{s_1}(x') - f_{s_1}(x)}{x' - x} = \text{rslope}(f_{s_1}, x)$.

2. If $f_{s_1}(x) = f_{s_2}(x)$ and $\text{lslope}(f_{s_1}, x) = \text{lslope}(f_{s_2}, x)$ then for a sequence $(x_i)_{i \in \mathbb{N}}$ that converges to x from the left, $\lim_{i \rightarrow \infty} \frac{f_{s_0}(x_i + \boldsymbol{\varrho}_1(s_0)) - f_{s_0}(x + \boldsymbol{\varrho}_1(s_0))}{x_i - x} = \text{lslope}(f_{s_0}, x + \boldsymbol{\varrho}_1(s_0))$. By the characterisation of the Pareto curve, $f_{s_0}(x_i + \boldsymbol{\varrho}_1(s_0)) \in \{\boldsymbol{\varrho}_2(s_0) + f_{s_1}(x_i), \boldsymbol{\varrho}_2(s_0) + f_{s_2}(x_i)\}$, so there are infinitely many x_i in the sequence for which $\boldsymbol{\varrho}_2(s_0) + f_{s_1}(x_i) = f_{s_0}(x_i + \boldsymbol{\varrho}_1(s_0))$ or there are infinitely many x_i in the sequence for which $\boldsymbol{\varrho}_2(s_0) + f_{s_2}(x_i) = f_{s_0}(x_i + \boldsymbol{\varrho}_1(s_0))$. Without loss of generality we assume it is for s_1 . $\text{lslope}(f_{s_1}, x) = \lim_{i \rightarrow \infty} \frac{f_{s_1}(x_i) - f_{s_1}(x)}{x_i - x} = \lim_{i \rightarrow \infty} \frac{f_{s_0}(x_i + \boldsymbol{\varrho}_1(s_0)) - f_{s_0}(x + \boldsymbol{\varrho}_1(s_0))}{x_i - x} = \text{lslope}(f_{s_0}, x + \boldsymbol{\varrho}_1(s_0))$. Hence, we proved that $\text{lslope}(f_{s_0}, x + \boldsymbol{\varrho}_1(s_0)) \in \{\boldsymbol{\varrho}_2(s_0) + \text{lslope}(f_{s_1}, x), \boldsymbol{\varrho}_2(s_0) + \text{lslope}(f_{s_2}, x)\}$. We could prove similarly, by considering a sequence x_i that converges from the right, that $\text{rslope}(f_{s_0}, x + \boldsymbol{\varrho}_1(s_0)) \in \{\boldsymbol{\varrho}_2(s_0) + \text{rslope}(f_{s_1}, x), \boldsymbol{\varrho}_2(s_0) + \text{rslope}(f_{s_2}, x)\}$.

We will now show that if $f_{s_1}(x) = f_{s_2}(x)$ and $\text{lslope}(f_{s_1}, x) < \text{lslope}(f_{s_2}, x)$, then $\text{lslope}(f_{s_0}, x + \boldsymbol{\varrho}_1(s_0)) = \text{lslope}(f_{s_2}, x)$, which shows the property. Since:

$$\lim_{x' \rightarrow x^-} \frac{f_{s_1}(x') - f_{s_1}(x)}{x' - x} = \text{lslope}(f_{s_1}, x) < \text{lslope}(f_{s_2}, x) = \lim_{x' \rightarrow x^-} \frac{f_{s_2}(x') - f_{s_2}(x)}{x' - x},$$

there is $\varepsilon > 0$, such that for all $x' \in [x - \varepsilon, x)$, $\frac{f_{s_1}(x') - f_{s_1}(x)}{x' - x} < \frac{f_{s_2}(x') - f_{s_2}(x)}{x' - x}$. This implies that $f_{s_1}(x') > f_{s_2}(x')$, because $x' - x < 0$ and therefore $f_{s_0}(x' + \boldsymbol{\varrho}_1(s_0)) = \boldsymbol{\varrho}_2(s_0) + f_{s_2}(x')$ (by characterization of the Pareto curve). Hence,

$$\begin{aligned} \text{lslope}(f_{s_0}, x + \boldsymbol{\varrho}_1(s_0)) &= \lim_{x' \rightarrow x^-} \frac{f_{s_0}(x' + \boldsymbol{\varrho}_1(s_0)) - f_{s_0}(x + \boldsymbol{\varrho}_1(s_0))}{x' - x} \\ &= \lim_{x' \rightarrow x^-} \frac{f_{s_2}(x') - f_{s_2}(x)}{x' - x} = \text{lslope}(f_{s_2}, x). \end{aligned}$$

The proof is quite similar for the right slope. Assume $\text{rslope}(f_{s_1}, x) < \text{rslope}(f_{s_2}, x)$. Since:

$$\begin{aligned} \lim_{x' \rightarrow x^+} \frac{f_{s_1}(x') - f_{s_1}(x)}{x' - x} &= \text{rslope}(f_{s_1}, x) < \text{rslope}(f_{s_2}, x) \\ &= \lim_{x' \rightarrow x^+} \frac{f_{s_2}(x') - f_{s_2}(x)}{x' - x}, \end{aligned}$$

there is $\varepsilon > 0$, such that for all $x' \in (x, x + \varepsilon)$, $\frac{f_{s_1}(x') - f_{s_1}(x)}{x' - x} < \frac{f_{s_2}(x') - f_{s_2}(x)}{x' - x}$. This implies that $f_{s_1}(x') < f_{s_2}(x')$, and therefore $f_{s_0}(x' + \boldsymbol{\varrho}_1(s_0)) = \boldsymbol{\varrho}_2(s_0) + f_{s_1}(x')$

(by characterisation of the Pareto curve). Hence,

$$\begin{aligned}
\text{rslope}(f_{s_0}, x + \boldsymbol{\varrho}_1(s_0)) &= \lim_{x' \rightarrow x^+} \frac{f_{s_0}(x' + \boldsymbol{\varrho}_1(s_0)) - f_{s_0}(x + \boldsymbol{\varrho}_1(s_0))}{x' - x} \\
&= \lim_{x' \rightarrow x^+} \frac{f_{s_2}(x') - f_{s_2}(x)}{x' - x} \\
&= \text{rslope}(f_{s_2}, x).
\end{aligned}$$

□

Lemma 10. *Let s_0 be a Player 2 state with two successors s_1 and s_2 , and let \mathbf{p} be a left accumulation point of f_{s_0} . There is $\eta(s_0, \mathbf{p}) > 0$ such that for all $\varepsilon \in (0, \eta(s_0, \mathbf{p}))$, there is $s' \in \{s_1, s_2\}$, such that: 1. $\mathbf{p} - \boldsymbol{\varrho}_1(s_0)$ is a left accumulation point in $f_{s'}$; 2. $\text{lslope}(s_0, x) = \text{lslope}(s', \mathbf{p}_1 - \boldsymbol{\varrho}_1(s_0))$; 3. $f_{s_0}(\mathbf{p}_1 - \varepsilon) = f_{s'}(\mathbf{p}_1 - \varepsilon - \boldsymbol{\varrho}_1(s_0))$ and $\text{rslope}(f_{s_0}, \mathbf{p}_1 - \varepsilon) = \text{rslope}(f_{s'}, \mathbf{p}_1 - \varepsilon - \boldsymbol{\varrho}_1(s_0))$.*

Proof. We first show that for each x where f_{s_0} is defined, there is $s' \in \{s_1, s_2\}$ such that $\text{lslope}(s_0, x) = \text{lslope}(s', x - \boldsymbol{\varrho}_1(s_0))$. Let $(x_i)_{i \in \mathbb{N}}$ be a sequence converging towards x . By the characterization of the Pareto curve, for each $i \in \mathbb{N}$, $f_{s_0}(x_i) \in \{\boldsymbol{\varrho}_2(s_0) + f_{s_1}(x_i - \boldsymbol{\varrho}_1(s_0)), \boldsymbol{\varrho}_2(s_0) + f_{s_2}(x_i - \boldsymbol{\varrho}_1(s_0))\}$. It will be either f_{s_1} or f_{s_2} infinitely often. We write s' for a state such that $f_{s_0}(x_i) = \boldsymbol{\varrho}_2(s_0) + f_{s'}(x_i - \boldsymbol{\varrho}_1(s_0))$ infinitely often. We can extract a subsequence of x_i so that we can assume $f_{s_0}(x_i) = \boldsymbol{\varrho}_2(s_0) + f_{s'}(x_i - \boldsymbol{\varrho}_1(s_0))$ for all i . Since this sequence converges to x , by continuity of $f_{s'}$ (Lem. 3) and f_{s_0} , we have $\boldsymbol{\varrho}_2(s_0) + f_{s'}(x - \boldsymbol{\varrho}_1(s_0)) = f_{s_0}(x)$. Moreover: $\text{lslope}(s_0, x) = \lim_{x' \rightarrow x^-} \frac{f_{s_0}(x') - f_{s_0}(x)}{x' - x} = \lim_{i \rightarrow \infty} \frac{f_{s_0}(x_i) - f_{s_0}(x)}{x_i - x} = \lim_{i \rightarrow \infty} \frac{f_{s'}(x_i - \boldsymbol{\varrho}_1(s_0)) - f_{s'}(x - \boldsymbol{\varrho}_1(s_0))}{x_i - x} = \text{lslope}(s', x)$.

We now show we can extract $(x_{\tau(i)})_{i \in \mathbb{N}}$, an infinite subsequence of $(x_i)_{i \in \mathbb{N}}$, whose elements all have different slopes. Since \mathbf{p} is a left accumulation point for f_{s_0} , for each $x_{\tau(i)}$, there are two extremal points for f_{s_0} in $(x_{\tau(i)}, x)$. Because of Lem. 20.4, no more than two extremal points can have the same slope, the slope of second extremal point \mathbf{p}' is strictly less than that of $x_{\tau(i)}$. Thus, by choosing $\tau(i+1)$ such that $x_{\tau(i+1)} > \mathbf{p}'_1$, we ensure that the slope of all $x_{\tau(i)}$ is different. This shows we can extract an infinite subsequence of $(x_i)_{i \in \mathbb{N}}$ whose all elements have different slopes. To simplify notations, we will still write $(x_i)_{i \in \mathbb{N}}$ for this subsequence, and assume all slopes $\text{lslope}(f_{s_0}, x_i)$ are different.

Because we showed $\text{lslope}(s_0, x) \in \{\text{lslope}(s_1, x - \boldsymbol{\varrho}_1(s_0)), \text{lslope}(s_2, x - \boldsymbol{\varrho}_1(s_0))\}$ for all points x where f_{s_0} is defined, this is also the case for each x_i . We can extract an infinite subsequence of $(x_i)_{i \in \mathbb{N}}$ such that it always corresponds to the same state. We can then assume that there is $s' \in \{s_1, s_2\}$, such that for all $i \in \mathbb{N}$, $\text{lslope}(s_0, x_i) = \text{lslope}(s', x_i - \boldsymbol{\varrho}_1(s_0))$. Since the slopes of every x_i is different, by Lem. 20.5, this implies that there is an infinite number of extremal points in the neighbourhood of \mathbf{p} . This means that \mathbf{p} is a left accumulation point for either s_1 or s_2 .

We now prove point 3. If \mathbf{p} is not a left accumulation point for s_1 , let $x_1 = \sup\{x_1 < x \mid (x_1, f_{s_1}(x_1)) \text{ extremal in } f_{s_1}\}$. Since f_{s_1} and f_{s_2} are concave their curves can only intersect twice in $[x_1, x]$. Since f_{s_0} has an infinite number of

extremal points in this interval (shifted by $\mathbf{q}_1(s_0)$) and f_{s_1} has none, we can deduce from Lem. 20.5 and Lem. 23 that f_{s_2} is below f_{s_1} in a neighbourhood $[\mathbf{q}_1, x]$ of x . Hence, for all $x' \in [\mathbf{q}_1, x]$, $f_{s_0}(x') = \mathbf{q}_2(s_0) + f_{s_2}(x' - \mathbf{q}_1(s_0))$. So the property we want is satisfied by $s' = s_2$ and $\eta(s_0, \mathbf{p}) = x - \mathbf{q}_1$.

Similarly, if \mathbf{p} is not a left accumulation point for s_2 , then $s' = s_1$ and some $\eta(s_0, \mathbf{p})$ witness the property.

If \mathbf{p} is a left accumulation point for s_1 but $\text{lslope}(f_{s_1}, \mathbf{p}_1 - \mathbf{q}_1(s_0)) \neq \text{lslope}(f_{s_0}, \mathbf{p}_1)$, then we can conclude from Lem. 23 that $\text{lslope}(f_{s_1}, \mathbf{p}_1) < \text{lslope}(f_{s_0}, \mathbf{p}_1) = \text{lslope}(f_{s_2}, \mathbf{p}_1)$. Hence, on some left neighbourhood of x , f_{s_2} is strictly below f_{s_1} . This means that the curves f_{s_0} and f_{s_2} (shifted by $\mathbf{q}_1(s_0)$) coincide on some neighbourhood of x . So the property we want is satisfied by $s' = s_2$ and some $\eta(s_0, \mathbf{p})$.

Similarly, if $\text{lslope}(f_{s_1}, \mathbf{p}_1) \neq \text{lslope}(f_{s_0}, \mathbf{p}_1)$, then $s' = s_1$ and some $\eta(s_0, \mathbf{p})$ witnesses the property.

In the other cases, both s_1 satisfies point (1) and (2). By Lem. 23, there is $s' \in \{s_1, s_2\}$, such that $\text{rslope}(f_{s'}, x' - \mathbf{q}_1(s_0)) \leq \text{rslope}(f_{s_0}, x')$. So the property is satisfied for s' and any $\eta(s_0, \mathbf{p})$ such that $x - \eta(s_0, \mathbf{p})$ is still in the domain of f_{s_0} . \square

C Evolution of the slope in stochastic states (proof of Lem. 11)

Lemma 25. *Let s_0 be a stochastic state with two successors s_1 and s_2 . If \mathbf{p} is a point of f_{s_0} then there are $\mathbf{q} \in f_{s_1}$ and $\mathbf{r} \in f_{s_2}$ such that $\mathbf{p} = \Delta(s_0, s_1) \cdot \mathbf{q} + \Delta(s_0, s_2) \cdot \mathbf{r}$. Moreover if \mathbf{p}' is another point of f_{s_0} with $\mathbf{p}_1 \leq \mathbf{p}'_1$, then we can chose $\mathbf{q}, \mathbf{q}' \in f_{s_1}$, $\mathbf{r}, \mathbf{r}' \in f_{s_2}$ such that $\mathbf{p} = \Delta(s_0, s_1) \cdot \mathbf{q} + \Delta(s_0, s_2) \cdot \mathbf{r}$, $\mathbf{p}' = \Delta(s_0, s_1) \cdot \mathbf{q}' + \Delta(s_0, s_2) \cdot \mathbf{r}'$, $\mathbf{q}_1 \leq \mathbf{q}'_1$ and $\mathbf{r}_1 \leq \mathbf{r}'_1$.*

Proof. Because \mathbf{p} is achievable in s_0 , by characterization of the Pareto curve, there are \mathbf{q} achievable in s_1 and \mathbf{r} achievable in s_2 such that $\mathbf{p} = \Delta(s_0, s_1) \cdot \mathbf{q} + \Delta(s_0, s_2) \cdot \mathbf{r}$. We assume towards a contradiction that \mathbf{q} is not from f_{s_1} (i.e. it lies strictly below the Pareto curve); the proof works in the same way for \mathbf{r} in s_2 . Then there would be some $\mathbf{q}' > \mathbf{q}$ that is achievable in s_1 . By characterization of the Pareto curve $\mathbf{p}' = \Delta(s_0, s_1) \cdot \mathbf{q}' + \Delta(s_0, s_2) \cdot \mathbf{r}$ is also achievable. However, we have $\mathbf{p}' > \mathbf{p}$ which is a contradiction with the fact that $\mathbf{p} \in f_{s_0}$. This shows the first part of the lemma.

Now for the second part of the lemma, assume that $\mathbf{p}_1 \leq \mathbf{p}'_1$ and $\mathbf{q}_1 > \mathbf{q}'_1$; the proof would work the same way if $\mathbf{r}_1 > \mathbf{r}'_1$. Then we must have $\mathbf{r}_1 < \mathbf{r}'_1$ since $\Delta(s_0, s_1) \cdot \mathbf{q}_1 + \Delta(s_0, s_2) \cdot \mathbf{r}_1 = \mathbf{p}_1 \leq \mathbf{p}'_1 = \Delta(s_0, s_1) \cdot \mathbf{q}'_1 + \Delta(s_0, s_2) \cdot \mathbf{r}'_1$. Let us write $m(\mathbf{q}, \mathbf{r}) = \Delta(s_0, s_1) \cdot \mathbf{q} + \Delta(s_0, s_2) \cdot \mathbf{r}$ to simplify notation. By monotonicity of m , $m(\mathbf{q}_1, \mathbf{r}'_1) > m(\mathbf{q}'_1, \mathbf{r}'_1) = \mathbf{p}'_1 \geq \mathbf{p}_1 = m(\mathbf{q}_1, \mathbf{r}_1)$. By continuity of m , there is $x' \in [\mathbf{r}_1, \mathbf{r}'_1]$ such that $m(\mathbf{q}_1, x') = \mathbf{p}'_1$ and similarly $x \in [\mathbf{r}_1, \mathbf{r}'_1]$ such that $m(\mathbf{q}'_1, x) = \mathbf{p}_1$.

Let us show $x - \mathbf{r}_1 = \mathbf{r}'_1 - x'$.

$$\begin{aligned} \Delta(s_0, s_1) \cdot \mathbf{q}'_1 + \Delta(s_0, s_2) \cdot x &= m(\mathbf{q}_1, x') = \mathbf{p}_1 = \Delta(s_0, s_1) \cdot \mathbf{q}_1 + \Delta(s_0, s_2) \cdot \mathbf{r}_1 \\ \Delta(s_0, s_2) \cdot (x - \mathbf{r}_1) &= \Delta(s_0, s_1) \cdot (\mathbf{q}_1 - \mathbf{q}'_1) \end{aligned}$$

Similarly:

$$\begin{aligned}\Delta(s_0, s_1) \cdot \mathbf{q}_1 + \Delta(s_0, s_2) \cdot x' &= \mathbf{p}'_1 = \Delta(s_0, s_1) \cdot \mathbf{q}'_1 + \Delta(s_0, s_2) \cdot \mathbf{r}'_1 \\ \Delta(s_0, s_2) \cdot (x' - \mathbf{r}'_1) &= \Delta(s_0, s_1) \cdot (\mathbf{q}'_1 - \mathbf{q}_1)\end{aligned}$$

Let us write $\alpha = \frac{\Delta(s_0, s_1)}{\Delta(s_0, s_2)} \cdot (\mathbf{q}_1 - \mathbf{q}'_1) = x - \mathbf{r}_1 = \mathbf{r}'_1 - x'$. By concavity of f_{s_2} :

$$\begin{aligned}f_{s_2}(x') &\geq f_{s_2}(\mathbf{r}'_1) + (x' - \mathbf{r}'_1) \cdot \frac{f_{s_2}(\mathbf{r}_1) - f_{s_2}(\mathbf{r}'_1)}{\mathbf{r}_1 - \mathbf{r}'_1} \\ f_{s_2}(x') - f_{s_2}(\mathbf{r}'_1) &\geq -\alpha \cdot \frac{f_{s_2}(\mathbf{r}_1) - f_{s_2}(\mathbf{r}'_1)}{\mathbf{r}_1 - \mathbf{r}'_1} \\ m(f_{s_1}(\mathbf{q}_1) - f_{s_1}(\mathbf{q}'_1), f_{s_2}(x') - f_{s_2}(\mathbf{r}'_1)) &\geq m\left(f_{s_1}(\mathbf{q}_1) - f_{s_1}(\mathbf{q}'_1), -\alpha \cdot \frac{f_{s_2}(\mathbf{r}_1) - f_{s_2}(\mathbf{r}'_1)}{\mathbf{r}_1 - \mathbf{r}'_1}\right)\end{aligned}$$

Similarly:

$$\begin{aligned}f_{s_2}(x) &\geq f_{s_2}(\mathbf{r}_1) + (x - \mathbf{r}_1) \cdot \frac{f_{s_2}(\mathbf{r}_1) - f_{s_2}(\mathbf{r}'_1)}{\mathbf{r}_1 - \mathbf{r}'_1} \\ f_{s_2}(x) - f_{s_2}(\mathbf{r}_1) &\geq \alpha \cdot \frac{f_{s_2}(\mathbf{r}_1) - f_{s_2}(\mathbf{r}'_1)}{\mathbf{r}_1 - \mathbf{r}'_1} \\ m(f_{s_1}(\mathbf{q}'_1) - f_{s_1}(\mathbf{q}_1), f_{s_2}(x) - f_{s_2}(\mathbf{r}_1)) &\geq m\left(f_{s_1}(\mathbf{q}'_1) - f_{s_1}(\mathbf{q}_1), \alpha \cdot \frac{f_{s_2}(\mathbf{r}_1) - f_{s_2}(\mathbf{r}'_1)}{\mathbf{r}_1 - \mathbf{r}'_1}\right) \\ &\geq -m\left(f_{s_1}(\mathbf{q}_1) - f_{s_1}(\mathbf{q}'_1), -\alpha \cdot \frac{f_{s_2}(\mathbf{r}_1) - f_{s_2}(\mathbf{r}'_1)}{\mathbf{r}_1 - \mathbf{r}'_1}\right)\end{aligned}$$

Hence, either $m(f_{s_1}(\mathbf{q}_1) - f_{s_1}(\mathbf{q}'_1), f_{s_2}(x') - f_{s_2}(\mathbf{r}'_1)) \geq 0$ or $m(f_{s_1}(\mathbf{q}'_1) - f_{s_1}(\mathbf{q}_1), f_{s_2}(x) - f_{s_2}(\mathbf{r}_1)) \geq 0$. In the first case $m(f_{s_1}(\mathbf{q}_1), f_{s_2}(x')) \geq m(f_{s_1}(\mathbf{q}'_1), f_{s_2}(\mathbf{r}'_1)) = \mathbf{p}'_2$. So we could have chosen \mathbf{q} and $\mathbf{r}'' = (x', f_{s_2}(x'))$ instead of \mathbf{q}' and \mathbf{r}' respectively, and all the properties required in the lemma are satisfied: $\mathbf{p} = m(\mathbf{q}, \mathbf{r})$, $\mathbf{p}' = m(\mathbf{q}, \mathbf{r}'')$, $\mathbf{q}_1 \leq \mathbf{q}_1$ and $\mathbf{r}_1 \leq \mathbf{r}''_1$.

In the second case $m(f_{s_1}(\mathbf{q}'_1), f_{s_2}(x)) \geq m(f_{s_1}(\mathbf{q}_1), f_{s_2}(\mathbf{r}_1)) = \mathbf{p}_2$. So we could have chosen \mathbf{q}' and $\mathbf{r}'' = (x, f_{s_2}(x))$ instead of \mathbf{q} and \mathbf{r} respectively, and all the properties required in the lemma are satisfied: $\mathbf{p} = m(\mathbf{q}', \mathbf{r}'')$, $\mathbf{p}' = m(\mathbf{q}', \mathbf{r}')$, $\mathbf{q}'_1 \leq \mathbf{q}'_1$ and $\mathbf{r}''_1 \leq \mathbf{r}'_1$. \square

Lemma 26. *Let s_0 be a stochastic state with two successors s_1 and s_2 . If \mathbf{p} is a point of f_{s_0} and f_{s_1} has a point \mathbf{q} and f_{s_2} a point \mathbf{r} such that $\mathbf{p} = \Delta(s_0, s_1) \cdot \mathbf{q} + \Delta(s_0, s_2) \cdot \mathbf{r}$, then if $\text{lslope}(s_1, \mathbf{q}_1)$ and $\text{lslope}(s_2, \mathbf{r}_1)$ are defined then $\text{lslope}(s_0, \mathbf{p}_1) = \min(\text{lslope}(s_1, \mathbf{q}_1), \text{lslope}(s_2, \mathbf{r}_1))$.*

Proof. We first show that $\text{lslope}(s_0, \mathbf{p}_1) \leq \text{lslope}(s_1, \mathbf{q}_1)$. We have that for all ε : $f_{s_0}(\mathbf{p}_1 - \varepsilon) \geq \Delta(s_0, s_1) \cdot f_{s_1}(\mathbf{q}_1 - \frac{\varepsilon}{\Delta(s_0, s_1)}) + \Delta(s_0, s_2) \cdot f_{s_2}(\mathbf{r}_1)$ because of the characterisation of the Pareto curve and the fact that $\mathbf{p}_1 - \varepsilon = \Delta(s_0, s_1) \cdot (\mathbf{q}_1 -$

$\frac{\varepsilon}{\Delta(s_0, s_1)}) + \Delta(s_0, s_2) \cdot \mathbf{r}_1$. Therefore:

$$\begin{aligned}
\lim_{\varepsilon \rightarrow 0^+} \frac{f_{s_0}(\mathbf{p}_1) - f_{s_0}(\mathbf{p}_1 - \varepsilon)}{\varepsilon} &= \lim_{\varepsilon \rightarrow 0^+} \frac{\Delta(s_0, s_1) \cdot f_{s_1}(\mathbf{q}_1) + \Delta(s_0, s_2) \cdot f_{s_2}(\mathbf{r}_1) - f_{s_0}(\mathbf{p}_1 - \varepsilon)}{\varepsilon} \\
&\leq \lim_{\varepsilon \rightarrow 0^+} \frac{\Delta(s_0, s_1) \cdot (f_{s_1}(\mathbf{q}_1) - f_{s_1}(\mathbf{q}_1 - \frac{\varepsilon}{\Delta(s_0, s_1)}))}{\varepsilon} \\
&\leq \lim_{\varepsilon \rightarrow 0^+} \frac{(f_{s_1}(\mathbf{q}_1) - f_{s_1}(\mathbf{q}_1 - \frac{\varepsilon}{\Delta(s_0, s_1)}))}{\frac{\varepsilon}{\Delta(s_0, s_1)}} \\
&\leq \lim_{\varepsilon \rightarrow 0^+} \frac{f_{s_1}(\mathbf{q}_1) - f_{s_1}(\mathbf{q}_1 - \varepsilon)}{\varepsilon}
\end{aligned}$$

Therefore, $\text{lslope}(s_0, \mathbf{p}_1) \leq \text{lslope}(s_1, \mathbf{q}_1)$. The proof also works for \mathbf{r} and therefore $\text{lslope}(s_0, \mathbf{p}_1) \leq \min(\text{lslope}(s_1, \mathbf{q}_1), \text{lslope}(s_2, \mathbf{r}_1))$.

We now show that $\text{lslope}(s_0, \mathbf{p}_1) \geq \min(\text{lslope}(s_1, \mathbf{q}_1), \text{lslope}(s_2, \mathbf{r}_1))$. Assume towards a contradiction that $\text{lslope}(s_0, \mathbf{p}_1) < \min(\text{lslope}(s_1, \mathbf{q}_1), \text{lslope}(s_2, \mathbf{r}_1))$. By definition of lslope there exists $x' < \mathbf{p}_1$ such that:

$$\frac{f_{s_0}(x') - f_{s_0}(\mathbf{p}_1)}{x' - \mathbf{p}_1} < \min(\text{lslope}(s_1, \mathbf{q}_1), \text{lslope}(s_2, \mathbf{r}_1))$$

By Lem. 25, there are $\mathbf{q}' \in f_{s_1}$ and $\mathbf{r}' \in f_{s_2}$ such that $(x', f_{s_0}(x')) = \Delta(s_0, s_1) \cdot \mathbf{q}' + \Delta(s_0, s_2) \cdot \mathbf{r}'$ and $\mathbf{q}'_1 < \mathbf{q}_1, \mathbf{r}'_1 < \mathbf{r}_1$. This gives:

$$\begin{aligned}
\min(\text{lslope}(s_1, \mathbf{q}_1), \text{lslope}(s_2, \mathbf{r}_1)) &> \frac{f_{s_0}(x') - f_{s_0}(\mathbf{p}_1)}{x' - \mathbf{p}_1} \\
&> \frac{\Delta(s_0, s_1) \cdot \mathbf{q}'_2 + \Delta(s_0, s_2) \cdot \mathbf{r}'_2 - \Delta(s_0, s_1) \cdot \mathbf{q}_2 - \Delta(s_0, s_2) \cdot \mathbf{r}_2}{\Delta(s_0, s_1) \cdot \mathbf{q}'_1 + \Delta(s_0, s_2) \cdot \mathbf{r}'_1 - \Delta(s_0, s_1) \cdot \mathbf{q}_1 - \Delta(s_0, s_2) \cdot \mathbf{r}_1} \\
&> \frac{\Delta(s_0, s_1) \cdot (\mathbf{q}'_2 - \mathbf{q}_2) + \Delta(s_0, s_2) \cdot (\mathbf{r}'_2 - \mathbf{r}_2)}{\Delta(s_0, s_1) \cdot (\mathbf{q}'_1 - \mathbf{q}_1) + \Delta(s_0, s_2) \cdot (\mathbf{r}'_1 - \mathbf{r}_1)}
\end{aligned}$$

By concavity of the Pareto curves:

$$\begin{aligned}
\frac{\mathbf{q}'_2 - \mathbf{q}_2}{\mathbf{q}'_1 - \mathbf{q}_1} &\geq \text{lslope}(s_1, \mathbf{q}_1) \\
\mathbf{q}'_2 - \mathbf{q}_2 &\leq \text{lslope}(s_1, \mathbf{q}_1) \cdot (\mathbf{q}'_1 - \mathbf{q}_1) \quad \text{because } \mathbf{q}'_1 - \mathbf{q}_1 \text{ is negative} \\
\frac{\mathbf{r}'_2 - \mathbf{r}_2}{\mathbf{r}'_1 - \mathbf{r}_1} &\geq \text{lslope}(s_2, \mathbf{r}_1) \\
\mathbf{r}'_2 - \mathbf{r}_2 &\leq \text{lslope}(s_2, \mathbf{r}_1) \cdot (\mathbf{r}'_1 - \mathbf{r}_1)
\end{aligned}$$

And since $\Delta(s_0, s_1) \cdot (\mathbf{q}'_1 - \mathbf{q}_1) + \Delta(s_0, s_2) \cdot (\mathbf{r}'_1 - \mathbf{r}_1)$ is negative:

$$\begin{aligned}
& \frac{\Delta(s_0, s_1) \cdot (\mathbf{q}'_2 - \mathbf{q}_2) + \Delta(s_0, s_2) \cdot (\mathbf{r}'_2 - \mathbf{r}_2)}{\Delta(s_0, s_1) \cdot (\mathbf{q}'_1 - \mathbf{q}_1) + \Delta(s_0, s_2) \cdot (\mathbf{r}'_1 - \mathbf{r}_1)} \\
& \geq \frac{\Delta(s_0, s_1) \cdot \text{lslope}(s_1, \mathbf{q}_1) \cdot (\mathbf{q}'_1 - \mathbf{q}_1) + \Delta(s_0, s_2) \cdot \text{lslope}(s_2, \mathbf{r}_1) \cdot (\mathbf{r}'_2 - \mathbf{r}_2)}{\Delta(s_0, s_1) \cdot (\mathbf{q}'_1 - \mathbf{q}_1) + \Delta(s_0, s_2) \cdot (\mathbf{r}'_1 - \mathbf{r}_1)} \\
& \geq \frac{\Delta(s_0, s_1) \cdot \min(\text{lslope}(s_1, \mathbf{q}_1), \text{lslope}(s_2, \mathbf{r}_1)) \cdot (\mathbf{q}'_1 - \mathbf{q}_1)}{\Delta(s_0, s_1) \cdot (\mathbf{q}'_1 - \mathbf{q}_1) + \Delta(s_0, s_2) \cdot (\mathbf{r}'_1 - \mathbf{r}_1)} \\
& \quad + \frac{\Delta(s_0, s_2) \cdot \min(\text{lslope}(s_1, \mathbf{q}_1), \text{lslope}(s_2, \mathbf{r}_1)) \cdot (\mathbf{r}'_2 - \mathbf{r}_2)}{\Delta(s_0, s_1) \cdot (\mathbf{q}'_1 - \mathbf{q}_1) + \Delta(s_0, s_2) \cdot (\mathbf{r}'_1 - \mathbf{r}_1)} \\
& \geq \min(\text{lslope}(s_1, \mathbf{q}_1), \text{lslope}(s_2, \mathbf{r}_1))
\end{aligned}$$

This would imply $\min(\text{lslope}(s_1, \mathbf{q}_1), \text{lslope}(s_2, \mathbf{r}_1)) > \min(\text{lslope}(s_1, \mathbf{q}_1), \text{lslope}(s_2, \mathbf{r}_1))$, hence a contradiction. Therefore, $\text{lslope}(s_0, \mathbf{p}_1) \geq \min(\text{lslope}(s_1, \mathbf{q}_1), \text{lslope}(s_2, \mathbf{r}_1))$, which concludes the proof. \square

Lemma 27. *Let s_0 be a stochastic state with two successors s_1 and s_2 . If \mathbf{p} is an extremal point in s_0 then s_1 has an extremal point \mathbf{q} and s_2 has an extremal point \mathbf{r} such that $\mathbf{p} = \Delta(s_0, s_1) \cdot \mathbf{q} + \Delta(s_0, s_2) \cdot \mathbf{r}$. Moreover \mathbf{q} and \mathbf{r} are uniquely defined and if $\text{lslope}(s_1, \mathbf{q}_1)$ and $\text{lslope}(s_2, \mathbf{r}_1)$ are defined then $\text{lslope}(s_0, \mathbf{p}_1) = \min(\text{lslope}(s_1, \mathbf{q}_1), \text{lslope}(s_2, \mathbf{r}_1))$. If only one of the two slopes is defined then $\text{lslope}(s_0, \mathbf{p}_1)$ is equal to that slope, and if none of the two is defined then $\text{lslope}(s_0, \mathbf{p}_1)$ is undefined.*

Proof. By the fixpoint characterisation of the achievable points, $\mathbf{p} = \Delta(s_0, s_1) \cdot \mathbf{q} + \Delta(s_0, s_2) \cdot \mathbf{r}$ for some \mathbf{r} achievable in s_2 and \mathbf{q} achievable in s_1 .

We first show that \mathbf{q} is extremal in s_1 and \mathbf{r} is extremal in s_2 . Assume towards a contradiction that there are $\mathbf{q}^1, \mathbf{q}^2$ achievable from s_1 such that $\mathbf{q} \in \text{conv}(\mathbf{q}^1, \mathbf{q}^2)$, then $\mathbf{p} \in \text{conv}(\Delta(s_0, s_1) \cdot \mathbf{q}^1 + \Delta(s_0, s_2) \cdot \mathbf{r}, \Delta(s_0, s_1) \cdot \mathbf{q}^2 + \Delta(s_0, s_2) \cdot \mathbf{r})$ and both these points are achievable from s_0 because of the characterization of the Pareto curve. This contradicts that \mathbf{p} is an extremal point. The proof works similarly for \mathbf{r} , and therefore \mathbf{q} and \mathbf{r} are extremal in f_{s_1} and f_{s_2} respectively.

We now prove that \mathbf{q} and \mathbf{r} are uniquely defined. Assume towards a contradiction that there are $(\mathbf{q}^1, \mathbf{r}^1) \neq (\mathbf{q}^2, \mathbf{r}^2)$ with \mathbf{q}^1 and \mathbf{q}^2 extremal in s_1 and \mathbf{r}^1 and \mathbf{r}^2 extremal in s_2 such that $\mathbf{p} = \Delta(s_0, s_1) \cdot \mathbf{q}^1 + \Delta(s_0, s_2) \cdot \mathbf{r}^1$ and $\mathbf{p} = \Delta(s_0, s_1) \cdot \mathbf{q}^2 + \Delta(s_0, s_2) \cdot \mathbf{r}^2$. Note that $\mathbf{q}^1 \neq \mathbf{q}^2$ and $\mathbf{r}^1 \neq \mathbf{r}^2$. Then we would have $\mathbf{p} = \frac{1}{2} \cdot (\Delta(s_0, s_1) \cdot \mathbf{q}^1 + \Delta(s_0, s_2) \cdot \mathbf{r}^2) + \frac{1}{2} \cdot (\Delta(s_0, s_1) \cdot \mathbf{q}^2 + \Delta(s_0, s_2) \cdot \mathbf{r}^1)$. This means that $\mathbf{p} \in \text{conv}(\Delta(s_0, s_1) \cdot \mathbf{q}^1 + \Delta(s_0, s_2) \cdot \mathbf{r}^2, \Delta(s_0, s_1) \cdot \mathbf{q}^2 + \Delta(s_0, s_2) \cdot \mathbf{r}^1)$, and both these points are achievable, which contradicts the fact that \mathbf{p} is extremal. Therefore, \mathbf{q} and \mathbf{r} are uniquely defined.

If both the slopes of f_1 and f_2 are defined then Lem. 26 implies the desired property of the slope. Assume now that exactly one of them is not defined; without loss of generality we assume $\text{lslope}(f_1)$ is defined and $\text{lslope}(f_2)$ is not. Then this means that f_2 is not defined at the left of \mathbf{r}_1 . Let $x' < \mathbf{p}_1$ such that $f_0(x')$ is defined. By Lem. 26, there are \mathbf{q}', \mathbf{r}' such that $(x', f_0(x')) = \Delta(s_0, s_1) \cdot \mathbf{q}' + \Delta(s_0, s_2) \cdot \mathbf{r}'$ and $\mathbf{r}'_1 \leq \mathbf{r}_1$. Since f_2 is not defined at the left of

\mathbf{r}_1 , we have that $\mathbf{r}_1 = \mathbf{r}'_1$ and so $\mathbf{r}' = \mathbf{r}$. Therefore, at the left of x , $f_0(x') = \Delta(s_0, s_1) \cdot f_1\left(\mathbf{q}_1 + \frac{x' - x}{\Delta(s_0, s_1)}\right) + \Delta(s_0, s_2) \cdot \mathbf{r}$. Thus, $\text{lslope}(f_0, x) = \Delta(s_0, s_1) \cdot \frac{1}{\Delta(s_0, s_1)} \cdot \text{lslope}(f_1, \mathbf{q}_1) = \text{lslope}(f_1, \mathbf{q}_1)$.

Now in the case where none of the two slopes are defined, this means that f_1 and f_2 are not defined at the left of \mathbf{q}_1 and \mathbf{r}_1 respectively. By Lem. 26, we can deduce that there is no point of f_0 at the left of \mathbf{p} , hence $\text{lslope}(f_0, \mathbf{p}_1)$ is also not defined. \square

For the proof of Lem. 32 below we will need several properties of the dot product. These are formalised in the lemmas below.

Lemma 28. *Let $\mathbf{n} \in \mathbb{R}^2$ be a vector with $n_1 > 0$ and $n_2 > 0$. If \mathbf{p} maximises the dot product $\mathbf{n} \cdot \mathbf{p}$ among the achievable vectors \mathcal{A}_s , then $\mathbf{p} \in f_s$ (i.e. \mathbf{p} is a maximal point in \mathcal{A}_s).*

Proof. Assume towards a contradiction that there is $\mathbf{p}' \geq \mathbf{p}$ which is achievable. Either $\mathbf{p}'_1 > \mathbf{p}_1$ and $\mathbf{p}'_2 \geq \mathbf{p}_2$, or $\mathbf{p}'_1 \geq \mathbf{p}_1$ and $\mathbf{p}'_2 > \mathbf{p}_2$. Then as $n_1 > 0$ and $n_2 > 0$, $\mathbf{n} \cdot \mathbf{p}' > \mathbf{n} \cdot \mathbf{p}$. This contradicts that \mathbf{p} maximises the dot product. \square

Lemma 29. *Let $\mathbf{n} \in \mathbb{R}^2$ be a vector with $n_1 > 0$ and $n_2 > 0$. If \mathbf{p} maximises the dot product $\mathbf{n} \cdot \mathbf{p}$ among achievable vectors, then $\text{lslope}(f_s, \mathbf{p}_1) \geq \frac{-n_1}{n_2} \geq \text{rslope}(f_s, \mathbf{p}_1)$.*

Proof. Assume towards a contradiction that $\text{lslope}(f_s, \mathbf{p}_1) < \frac{-n_1}{n_2}$. Then there is $x' < \mathbf{p}_1$ such that:

$$\begin{aligned} \frac{f_s(x') - f_s(\mathbf{p}_1)}{x' - \mathbf{p}_1} &< \frac{-n_1}{n_2} \\ n_2 \cdot (f_s(x') - f_s(\mathbf{p}_1)) &> -n_1 \cdot (x' - \mathbf{p}_1) \quad \text{because } x' - \mathbf{p}_1 < 0 \\ n_1 \cdot x' + n_2 \cdot f_s(x') &> n_1 \cdot \mathbf{p}_1 + n_2 \cdot f_s(\mathbf{p}_1) \end{aligned}$$

This contradicts that \mathbf{p} maximises the dot product. \square

Lemma 30. *If $\mathbf{p} \in f_s$ then it maximises the dot product with $(-\text{lslope}(f_s, \mathbf{p}_1), 1)$ and with $(-\text{rslope}(f_s, \mathbf{p}_1), 1)$ among achievable vectors \mathcal{A}_s .*

Proof. Let \mathbf{p}' be a point of f_s . Since f_s is concave, \mathbf{p}' is below the line $\{(\mathbf{p}_1, f_s(\mathbf{p}_1)) + t \cdot (1, \text{lslope}(f_s, \mathbf{p}_1)) \mid t \in \mathbb{R}\}$. So $\mathbf{p}'_2 \leq f_s(\mathbf{p}_1) + (\mathbf{p}'_1 - \mathbf{p}_1) \cdot \text{lslope}(f_s, \mathbf{p}_1)$. Then, looking at the dot product:

$$\begin{aligned} (-\text{lslope}(f_s, \mathbf{p}_1), 1) \cdot \mathbf{p}' &= \mathbf{p}'_2 - \mathbf{p}'_1 \cdot \text{lslope}(f_s, \mathbf{p}_1) \\ &\leq f_s(\mathbf{p}_1) + (\mathbf{p}'_1 - \mathbf{p}_1) \cdot \text{lslope}(f_s, \mathbf{p}_1) - \mathbf{p}'_1 \cdot \text{lslope}(f_s, \mathbf{p}_1) \\ &\leq f_s(\mathbf{p}_1) - \mathbf{p}_1 \cdot \text{lslope}(f_s, \mathbf{p}_1) \\ &\leq f_s(\mathbf{p}_1) - \mathbf{p}_1 \cdot \text{lslope}(f_s, \mathbf{p}_1) \\ &\leq (-\text{lslope}(f_s, \mathbf{p}_1), 1) \cdot (\mathbf{p}_1, f_s(\mathbf{p}_1)) \\ &\leq (-\text{lslope}(f_s, \mathbf{p}_1), 1) \cdot \mathbf{p} \end{aligned}$$

The proof works similarly for $(-\text{rslope}(f_s, \mathbf{p}_1), 1)$. \square

Lemma 31. Let \mathbf{n} be a vector in \mathbb{R}^2 , $Y, Z \subseteq \mathbb{R}^2$, $\lambda_1, \lambda_2 \in \mathbb{R}_{>0}$, and $X = \lambda_1 Y + \lambda_2 Z$. Let $(\mathbf{y}, \mathbf{z}) \in Y \times Z$ and $\mathbf{x} = \lambda_1 \mathbf{y} + \lambda_2 \mathbf{z}$. Then \mathbf{y} and \mathbf{z} maximize the dot product with \mathbf{n} among vectors of Y and Z respectively, if and only if, \mathbf{x} maximizes the dot product with \mathbf{n} among vectors of X .

Proof. Assume that \mathbf{y} and \mathbf{z} maximise the dot product with \mathbf{n} and let $\mathbf{a} \in \lambda_1 Y + \lambda_2 Z$. We have that $\mathbf{a} \cdot \mathbf{n} = \lambda_1 \mathbf{b} \cdot \mathbf{n} + \lambda_2 \mathbf{c} \cdot \mathbf{n}$ for some $(\mathbf{b}, \mathbf{c}) \in Y \times Z$. This is smaller than $\lambda_1 \mathbf{y} \cdot \mathbf{n} + \lambda_2 \mathbf{z} \cdot \mathbf{n}$, since they maximize the dot product with \mathbf{n} within their respective sets. It is therefore smaller than $\mathbf{x} \cdot \mathbf{n}$.

Reciprocally, assume that \mathbf{x} maximises the dot product with \mathbf{n} . If there is $\mathbf{u} \in Y$ such that $\mathbf{u} \cdot \mathbf{n} > \mathbf{y} \cdot \mathbf{n}$ then $\lambda_1 \mathbf{u} + \lambda_2 \mathbf{z}$ has a greater dot product with \mathbf{n} than \mathbf{x} , which is a contradiction. \square

Lemma 32. Let s_0 be a stochastic state with two successors s_1 and s_2 . Let $\mathbf{p} \in f_{s_0}$, $\mathbf{q} \in f_{s_1}$, and $\mathbf{r} \in f_{s_2}$ such that $\mathbf{p} = \Delta(s_0, s_1) \cdot \mathbf{q} + \Delta(s_0, s_2) \cdot \mathbf{r}$. For all $\varepsilon > 0$, there are $\varepsilon_1, \varepsilon_2$ such that $\text{rslope}(f_{s_0}, \mathbf{p}_1 - \varepsilon) \geq \text{rslope}(f_{s_1}, \mathbf{q}_1 - \varepsilon_1)$, $\text{rslope}(f_{s_0}, \mathbf{p}_1 - \varepsilon) \geq \text{rslope}(f_{s_2}, \mathbf{r}_1 - \varepsilon_2)$, and $\varepsilon = \Delta(s_0, s_1) \cdot \varepsilon_1 + \Delta(s_0, s_2) \cdot \varepsilon_2$.

Proof. By Lem. 25, there are \mathbf{q}' and \mathbf{r}' such that $(\mathbf{p}_1 - \varepsilon, f_{s_0}(\mathbf{p}_1 - \varepsilon)) = \Delta(s_0, s_1) \cdot \mathbf{q}' + \Delta(s_0, s_2) \cdot \mathbf{r}'$ and $\text{lslope}(f_{s_0}, \mathbf{p}_1 - \varepsilon) = \min\{\text{lslope}(f_{s_1}, \mathbf{q}'_1), \text{lslope}(f_{s_2}, \mathbf{r}'_1)\}$. We let $\varepsilon_1 = \mathbf{r}_1 - \mathbf{r}'_1$ and $\varepsilon_2 = \mathbf{q}_1 - \mathbf{q}'_1$.

Let $\mathbf{n} = (-\text{rslope}(f_{s_0}, \mathbf{p}_1 - \varepsilon), 1)$ be a vector that follows the normal to the slope in $\mathbf{p}_1 - \varepsilon$. By Lem. 30, $\mathbf{p}' = (\mathbf{p}_1 - \varepsilon, f_{s_0}(\mathbf{p}_1 - \varepsilon))$ maximises the dot product with \mathbf{n} on the curve of f_{s_0} . By Lem. 31, it is also the case of \mathbf{q}' and \mathbf{r}' on the curve of f_{s_1} and f_{s_2} respectively. Therefore, by Lem. 29, $\text{lslope}(f_{s_1}, \mathbf{q}_1 - \varepsilon_1) \geq \text{rslope}(f_{s_0}, \mathbf{p}_1 - \varepsilon) \geq \text{rslope}(f_{s_1}, \mathbf{q}_1 - \varepsilon_1)$ and similarly $\text{rslope}(f_{s_0}, \mathbf{p}_1 - \varepsilon) \geq \text{rslope}(f_{s_2}, \mathbf{r}_1 - \varepsilon_2)$. \square

Intuitively the next lemma says that for a stochastic state s_0 with successors s_1 and s_2 , if s_1 has no left accumulation point, then the slopes decrease faster in s_2 than in s_0 .

Lemma 11. Let s_0 be a stochastic state with two successors s_1 and s_2 , and \mathbf{p} a left accumulation point of f_{s_0} . There are points \mathbf{q} and \mathbf{r} on f_{s_1} and f_{s_2} respectively such that $\mathbf{p} = \Delta(s_0, s_1) \cdot \mathbf{q} + \Delta(s_0, s_2) \cdot \mathbf{r}$. Moreover:

1. there is $(s', \mathbf{p}') \in \{(s_1, \mathbf{q}), (s_2, \mathbf{r})\}$ such that \mathbf{p}' is a left accumulation point of $f_{s'}$ and $\text{lslope}(f_{s_0}, \mathbf{p}_1) = \text{lslope}(f_{s'}, \mathbf{p}'_1)$;
2. there is $\eta(s_0, \mathbf{p}) > 0$ such that for all $\varepsilon \in [0, \eta(s_0, \mathbf{p}))$:
 - there are $\varepsilon_1, \varepsilon_2$ such that $\text{rslope}(f_{s_0}, \mathbf{p}_1 - \varepsilon) \geq \text{rslope}(f_{s_1}, \mathbf{q}_1 - \varepsilon_1)$, $\text{rslope}(f_{s_0}, \mathbf{p}_1 - \varepsilon) \geq \text{rslope}(f_{s_2}, \mathbf{r}_1 - \varepsilon_2)$, and $\varepsilon = \Delta(s_0, s_1) \cdot \varepsilon_1 + \Delta(s_0, s_2) \cdot \varepsilon_2$.
 - if \mathbf{r} is not a left accumulation point in f_{s_2} , or $\text{lslope}(f_{s_0}, \mathbf{p}_1) \neq \text{lslope}(f_{s_2}, \mathbf{r}_1)$, then $f_{s_0}(\mathbf{p}_1 - \varepsilon) = \Delta(s_0, s_1) \cdot f_{s_1}\left(\frac{\mathbf{p}_1 - \varepsilon - \Delta(s_0, s_2) \cdot \mathbf{r}_1}{\Delta(s_0, s_1)}\right) + \Delta(s_0, s_2) \cdot \mathbf{r}_2$.
 - symmetrically, if \mathbf{q} is not a left accumulation point in f_{s_1} , or $\text{lslope}(f_{s_0}, \mathbf{p}_1) \neq \text{lslope}(f_{s_1}, \mathbf{q}_1)$, then $f_{s_0}(\mathbf{p}_1 - \varepsilon) = \Delta(s_0, s_1) \cdot \mathbf{q}_2 + \Delta(s_0, s_2) \cdot f_{s_2}\left(\frac{\mathbf{p}_1 - \varepsilon - \Delta(s_0, s_1) \cdot \mathbf{q}_1}{\Delta(s_0, s_2)}\right)$.

Proof. Let $(\mathbf{p}^i)_{i \in \mathbb{N}}$ be a sequence of extremal points in s_0 with increasing first coordinate which converges towards \mathbf{p} . By Lem. 27, there are \mathbf{q}^i and \mathbf{r}^i extremal in f_{s_1} and f_{s_2} respectively, such that $\mathbf{p}^i = \Delta(s_0, s_1) \cdot \mathbf{q}^i + \Delta(s_0, s_2) \cdot \mathbf{r}^i$. Lem. 25 tells us that for a particular index i , we could choose $\mathbf{q}^i, \mathbf{r}^i, \mathbf{q}^{i+1}$ and \mathbf{r}^{i+1} such that $\mathbf{q}_1^i \leq \mathbf{q}_1^{i+1}$ and $\mathbf{r}_1^i \leq \mathbf{r}_1^{i+1}$. But since Lem. 27 shows that \mathbf{q}^i and \mathbf{r}^i are uniquely defined, the sequence indeed satisfies the fact that x -coordinates are increasing. The sequences \mathbf{q}^i and \mathbf{r}^i converge because their first coordinate are increasing and bounded. The limits \mathbf{q} and \mathbf{r} are such that $\mathbf{p} = \Delta(s_0, s_1) \cdot \mathbf{q} + \Delta(s_0, s_2) \cdot \mathbf{r}$.

1. Since \mathbf{p}^i contains an infinite number of different points, by Lem. 20.4, there should also be an infinite number of different slopes (no more than two extremal points can have the same slope). By Lem. 27, we have that for all index i , $\text{lslope}(s_0, \mathbf{p}^i) = \min(\text{lslope}(s_1, \mathbf{q}^i), \text{lslope}(s_2, \mathbf{r}^i))$. This means one of \mathbf{q}^i and \mathbf{r}^i gives an infinite number of slopes, and therefore also an infinite number of points. Let say that it is \mathbf{q}^i . Since the points of the sequence lie on the Pareto curve and converge to \mathbf{q} , \mathbf{q} is a left accumulation point. Moreover by Lem. 20.6, $\text{lslope}(f_{s_0}, \mathbf{p}) = \text{lslope}(f_{s_1}, \mathbf{q})$. Similarly, if the \mathbf{r}^i contains infinitely many different points, then \mathbf{r} is a left accumulation point and $\text{lslope}(f_{s_0}, \mathbf{p}) = \text{lslope}(f_{s_2}, \mathbf{r})$.

2. Note that $\text{lslope}(f_{s_0}, \cdot)$ is defined at a neighbourhood on the left of \mathbf{p}_1 because it is a left accumulation point. We let $\eta(s_0, \mathbf{p})$ be such that $\mathbf{p}_1 - \eta(s_0, \mathbf{p})$ is included in that neighbourhood. By Lem. 32, we have that for all $\varepsilon \in [0, \eta(s_0, \mathbf{p}))$, there are $\varepsilon_1, \varepsilon_2$ such that $\text{rslope}(f_{s_0}, \mathbf{p}_1 - \varepsilon) \geq \text{rslope}(f_{s_1}, \mathbf{q}_1 - \varepsilon_1)$, $\text{rslope}(f_{s_0}, \mathbf{p}_1 - \varepsilon) \geq \text{rslope}(f_{s_2}, \mathbf{r}_1 - \varepsilon_2)$, and $\varepsilon = \Delta(s_0, s_1) \cdot \varepsilon_1 + \Delta(s_0, s_2) \cdot \varepsilon_2$.

We now look at the cases where \mathbf{r} is not a left accumulation point of s_2 or $\text{lslope}(f_{s_0}, \mathbf{p}_1) \neq \text{lslope}(f_{s_2}, \mathbf{r}_1)$. In the first case, it is clear that the sequence of \mathbf{r}^i does not contain infinitely many different points. In the second case, by Lem. 27, $\text{lslope}(f_{s_0}, \mathbf{p}_1) < \text{lslope}(f_{s_2}, \mathbf{r}_1)$. By Lem. 20.6, the slope in \mathbf{p} is the limit of the slopes in the points \mathbf{p}^i . So there is an index j after which $\text{lslope}(f_{s_0}, \mathbf{p}_1^i) < \text{lslope}(f_{s_2}, \mathbf{r}_1)$ for $i \geq j$. By Lem. 30, \mathbf{r} maximises the dot product with $(-\text{lslope}(f_{s_2}, \mathbf{r}_1), 1)$ on the curve of f_{s_2} . By the same lemma, \mathbf{p} maximises the dot product with $(-\text{lslope}(f_{s_0}, \mathbf{p}_1), 1)$ on the curve of f_{s_0} . Since $\mathbf{p} = \Delta(s_0, s_1) \cdot \mathbf{q} + \Delta(s_0, s_2) \cdot \mathbf{r}$, Lem. 31 implies that \mathbf{r} also maximises the dot product with $(-\text{lslope}(f_{s_0}, \mathbf{p}_1), 1)$ on the curve of f_{s_2} . The point \mathbf{r} therefore maximises the dot product with all vectors $(\alpha, 1)$ with $\alpha \in [-\text{lslope}(f_{s_0}, \mathbf{p}_1), -\text{lslope}(f_{s_0}, \mathbf{p}_1)]$. This is in particular the case for $(-\text{lslope}(f_{s_0}, \mathbf{p}_1^i), 1)$ where $i \geq j$. By Lem. 31, this implies that $\mathbf{r}^i = \mathbf{r}$. Hence, in both cases, we can extract a subsequence of \mathbf{p}^i where \mathbf{r}^i is a constant \mathbf{r}^0 . Since the sequence \mathbf{r}^i converges to \mathbf{r} , we get $\mathbf{r}^0 = \mathbf{r}$. We have that $\mathbf{q}^i = \frac{\mathbf{p}^i + \Delta(s_0, s_2) \cdot \mathbf{r}}{\Delta(s_0, s_1)}$.

Let us show that for \mathbf{p}' on the Pareto curve of s_0 close enough at the left to \mathbf{p} there is \mathbf{q}' on the Pareto curve in s_1 such that $\mathbf{p}' = \Delta(s_0, s_1) \cdot \mathbf{q}' + \Delta(s_0, s_2) \cdot \mathbf{r}$. Let \mathbf{p}' such that $\mathbf{p}'_1 > \mathbf{p}_1^0$, we will show that such a \mathbf{q}' exists. By Lem. 25 there are \mathbf{q}^1 and \mathbf{r}^1 such that $\mathbf{p}' = \Delta(s_0, s_1) \cdot \mathbf{q}^1 + \Delta(s_0, s_2) \cdot \mathbf{r}^1$ and because $\mathbf{p}^0 = \Delta(s_0, s_1) \cdot \mathbf{q}^0 + \Delta(s_0, s_2) \cdot \mathbf{r}$, we can choose $\mathbf{r}'_1 \geq \mathbf{r}_1^0 = \mathbf{r}_1$. Since $\mathbf{p} = \Delta(s_0, s_1) \cdot \mathbf{q}^0 + \Delta(s_0, s_2) \cdot \mathbf{r}$, we can also choose \mathbf{q}^2 and \mathbf{r}^2 such that $\mathbf{p}' = \Delta(s_0, s_1) \cdot \mathbf{q}^2 + \Delta(s_0, s_2) \cdot \mathbf{r}^2$ and $\mathbf{r}_1^2 \leq \mathbf{r}_1^0 = \mathbf{r}_1$. There exists $\lambda_1, \lambda_2 \in [0, 1]$ whose sum is 1 and such that

$\lambda_1 \cdot \mathbf{r}^1 + \lambda_2 \cdot \mathbf{r}^2 = \mathbf{r}$. Moreover letting $\mathbf{q}' = \lambda_1 \cdot \mathbf{q}^1 + \lambda_2 \cdot \mathbf{q}^2$, we obtain that $\mathbf{p}' = \Delta(s_0, s_1) \cdot \mathbf{q}' + \Delta(s_0, s_2) \cdot \mathbf{r}$.

Therefore, for $x' \in [p_1^0, \mathbf{p}_1)$, $f_0(x') = \Delta(s_0, s_1) \cdot f_1(\mathbf{q}'_1) + \Delta(s_0, s_2) \cdot f_2(\mathbf{r}_1)$ where \mathbf{q}'_1 is such that $x' = \Delta(s_0, s_1) \cdot \mathbf{q}'_1 + \Delta(s_0, s_2) \cdot \mathbf{r}_1$. That means $f_0(x') = \Delta(s_0, s_1) \cdot f_1\left(\frac{x' + \Delta(s_0, s_2)}{\Delta(s_0, s_1)}\right) + \Delta(s_0, s_2) \cdot f_2(\mathbf{r}_1)$. \square

D Inverse betting game

D.1 Proof of Thm. 13

Theorem 13. *Let $\langle V, E, (v_0, c_0), w \rangle$ be a inverse betting game. Let $T \subseteq V$ be a target set and $B \in \mathbb{R}$ a bound. If from every vertex $v \in V$, Eve has a strategy to ensure visiting T then she has one to ensure visiting it with a valuation of the counter $c \geq 1$ or to exceed the bound, that is she can force a configuration in $(T \times [c_0, +\infty)) \cup (V \times [B, +\infty))$.*

Proof. Assuming Eve has a strategy to ensure visiting T , then she has a memoryless strategy to do so (see for example [10]). We write $\sigma: V \rightarrow V$ for the function on states associated to this memoryless strategy that ensures visiting T from v (it is easy to recover the full strategy from there: $h \mapsto \sigma(\text{last}(h))$). We also write $a(v)$ for the length of the longest path from v compatible with σ that does not reach T . Note that $a(v)$ is bounded by $|V|$ and decrease with each step compatible with σ .

We define a potential function over configurations: $p(v, c) = c + W^{a(v)} - W^{|V|}$. Note that because a is bounded, when p goes to infinity, c also goes to infinity.

The idea for our strategy is to never make this potential decrease. We show that it is possible to do so in each configuration that is not a target. Given a configuration (v, c) , let us write v_1 and v_2 the successors of v and c_1 and c_2 the respective valuations of these successors chosen by Adam. One of the successors is closer to T with respect to σ , so without loss of generality we assume that $a(v_1) \leq a(v) - 1$. We have that $c = w(v, v_1) \cdot c_1 + w(v, v_2) \cdot c_2$.

$$\begin{aligned}
p(v_1, c_1) &= c_1 + W^{a(v_1)} - W^{|V|} \\
&\geq c_1 + W^{a(v)-1} - W^{|V|} \quad (\text{as } a(v_1) \leq a(v) - 1 \text{ and } W \leq 1) \\
p(v_2, c_2) &= c_2 + W^{a(v_2)} - W^{|V|} \\
&\geq c_2 \quad (\text{as } a(v_2) \leq |V| \text{ and } W \leq 1) \\
w(v, v_1) \cdot p(v_1, c_1) + w(v, v_2) \cdot p(v_2, c_2) & \\
&\geq w(v, v_1) \cdot c_1 + w(v, v_1) \cdot (W^{a(v)-1} - W^{|V|}) + w(v, v_2) \cdot c_2 \\
&\geq c + w(v, v_1) \cdot (W^{a(v)-1} - W^{|V|}) \\
&\geq c + W \cdot (W^{a(v)-1} - W^{|V|}) \quad (\text{as } w(v, v_1) \geq W) \\
&\geq c + W^{a(v)} - W^{|V|+1} \\
&\geq p(v, c) + W^{|V|} - W^{|V|+1}
\end{aligned}$$

Since $w(v, v_1) + w(v, v_2) = 1$, either $p(v_1, c_1) \geq p(v, c) + W^{|V|} - W^{|V|+1}$ or $p(v_2, c_2) > p(v, c) + W^{|V|} - W^{|V|+1}$. We define σ' , to choose (v_1, c_1) in the first case and (v_2, c_2) in the second one. Along any path compatible with this strategy the potential at each step increases by at least $W^{|V|} - W^{|V|+1}$, which is strictly positive. This means that either it will reach a target (then $a(v)$ can no longer decrease) or it goes to infinity, and so does c .

D.2 Following a point close to the left accumulation point (proof of Lem. 16)

We consider a sequence of points that are $\theta(s_0, \mathbf{p}_0)$ close to \mathbf{p}_0 and with a slope that is decreasing at least as fast as that of their predecessors.

Lemma 16. *For stopping games, given $s_0, \mathbf{p}^0, \varepsilon_0$, such that $\varepsilon_0 < \theta(s_0, \mathbf{p}^0)$, there exists a finite sequence $\pi(s_0, \mathbf{p}^0, \varepsilon_0) = (s_i, \mathbf{p}^i, \varepsilon_i)_{i \leq j}$ such that:*

- $(s_i, \mathbf{p}^i)_{i \leq j}$ is a path in T_{s_0, \mathbf{p}^0} ;
- for all $i \leq j$, $\text{rslope}(f_{s_i}, p_1^i - \varepsilon_i) \geq \text{rslope}(f_{s_{i+1}}, p_1^{i+1} - \varepsilon_{i+1})$.
- either $s_j \in U_{s_0, \mathbf{p}^0}$ and $\varepsilon_j \geq \varepsilon_0$ or $\varepsilon_j \geq \theta(s_0, \mathbf{p}^0)$.

Proof. To construct this path, we will invoke results on inverse betting games presented in Sec. 4.2. Consider the inverse betting game given by T_{s_0, \mathbf{p}^0} in Sec. 4.2.

We show in every configuration $((s_i, \mathbf{p}^i), c_i)$ with $c_i \leq \theta(s_0, \mathbf{p}^0)$, Adam has a choice in its action such that the successor $((s_{i+1}, \mathbf{p}^{i+1}), c_{i+1})$ will be such that $\text{lslope}(f_{s_i}, p_1^i - c_i) \geq \text{lslope}(f_{s_{i+1}}, p_1^{i+1} - c_{i+1})$.

- For Player 1 states, this is thanks to Lem. 9: since $\varepsilon_i \leq \theta(s_0, \mathbf{p}^0) \leq \eta(s_i, \mathbf{p}^i)$, we have that there is s' in $\Delta(s_i)$ that is a successor of s_i in T_{s_0, \mathbf{p}^0} (because it has a left accumulation point \mathbf{p}^i and $\text{lslope}(s_i, p_1^i) = \text{lslope}(s', p_1^i)$) and such that $\text{rslope}(f_{s_i}, p_1^i - \varepsilon_i) \geq \text{rslope}(f_{s'}, p_1^i - \varepsilon_i)$. Since Adam controls the configurations corresponding to Player 1 states, he can chose the appropriate successor.
- For Player 2 states, thanks to Lem. 10: since $\varepsilon_i \leq \theta(s_0, \mathbf{p}^0) \leq \eta(s_i, \mathbf{p}^i)$, we have that there is s' in $\Delta(s_i)$ that is a successor of s_i in T_{s_0, \mathbf{p}^0} (because it has a left accumulation point \mathbf{p}^i and $\text{lslope}(s_i, p_1^i + \mathbf{q}_1(s_i)) = \text{lslope}(s', p_1^i)$) and such that $\text{rslope}(f_{s_i}, p_1^i + \mathbf{q}_1(s_i) - \varepsilon_i) = \text{rslope}(f_{s'}, p_1^i - \varepsilon_i)$. Since Adam controls the configurations corresponding to Player 2 states, he can chose the appropriate successor.
- If s_i is a stochastic state, then by Lem. 32 there are $c_1, c_2 \in \mathbb{R}$ such that the successors (s_1, \mathbf{q}) and (s_2, \mathbf{r}) of (s_i, \mathbf{p}^i) in T_{s_0, \mathbf{p}^0} , are such that $c_i = \Delta(s_0, s_i) \cdot c_1 + \Delta(s_i, s_2) \cdot c_2$ and $\text{rslope}(f_{s_i}, p_1^i - c_i) \geq \text{rslope}(f_{s_1}, \mathbf{q}_1 - c_1)$ and $\text{rslope}(f_{s_i}, p_1^i - c_i) \geq \text{rslope}(f_{s_2}, \mathbf{r}_1 - c_2)$. So by choosing c_1 for s_1 and c_2 for s_2 , Adam ensures that for all choices of Eve, $\text{rslope}(f_{s_i}, p_1^i - \varepsilon_i) \geq \text{rslope}(f_{s_{i+1}}, p_1^{i+1} - \varepsilon_{i+1})$.

With such choices for Adam, there is a strategy σ_\forall that ensures that $\text{rslope}(f_{s_i}, p_1^i - c_i)$ is decreasing along the outcome of the game.

By Cor. 15, for any bound B there is a strategy for **Eve** to ensure we reach a configuration in $(U_{s_0, \mathbf{p}^0} \times [c_0, +\infty)) \cup (V \times [B, +\infty))$. This is in particular the case for $B = \theta(s_0, \mathbf{p}^0)$, and we write σ_{\exists} the corresponding strategy.

The outcome ρ of $(\sigma_{\exists}, \sigma_{\forall})$ has both properties. We now distinguish two types of paths:

- If ρ reaches a configuration with credit greater than $\theta(s_0, \mathbf{p}^0)$, then let j be the first index where this happen. We have that for all $i < j$, $\text{rslope}(f_{s_i}, p_1^i - \varepsilon_i) \geq \text{rslope}(f_{s_{i+1}}, p_1^{i+1} - \varepsilon_{i+1})$, thanks to the construction of strategy σ_{\forall} .
- Otherwise, we have for all i that $\text{rslope}(f_{s_i}, p_1^i - \varepsilon_i) \geq \text{rslope}(f_{s_{i+1}}, p_1^{i+1} - \varepsilon_{i+1})$, thanks to the construction of strategy σ_{\forall} . Moreover since σ_{\exists} is winning and we do not get to a configuration in $V \times [\theta(s_0, \mathbf{p}^0), +\infty)$, ρ reaches U_{s_0, \mathbf{p}^0} with a credit greater than the initial credit that was ε_0 . Let j be the first index where this happens.

In both case we have that $\rho_{\leq j}$ is a witness of the property. \square

D.3 Proof of Lem. 17

Lemma 17. *For all states s with a left accumulation point \mathbf{p} and for all $\varepsilon < \theta(s, \mathbf{p})$, there is some (s', \mathbf{p}') reachable in $T_{s, \mathbf{p}}$ such that $\text{rslope}(f_{s'}, \mathbf{p}' - \theta(s_0, \mathbf{p}^0)) \leq \text{rslope}(f_s, \mathbf{p}_1 - \varepsilon)$.*

Proof. Consider the sequence $\pi(s, \mathbf{p}, \varepsilon)$ as defined in Lem. 16. Either for the last configuration, ε_j is greater than $\theta(s, \mathbf{p})$ in which case we directly get the property for $(s', \mathbf{p}') = (s_j, \mathbf{p}^j)$; or we reach U_{s_0, \mathbf{p}^0} . In this case, we have that $\varepsilon_j \geq \varepsilon$, and by Lem. 11.2, there is a successor $(s_{j+1}, \mathbf{p}^{j+1})$ in $T_{s, \mathbf{p}}$ such that for all $\varepsilon \leq \theta(s_0, \mathbf{p}^0)$, $f_{s_j}(p_1^j - \varepsilon) = \Delta(s_j, s_{j+1}) \cdot f_{s_{j+1}}\left(\frac{p_1^j - \varepsilon - \Delta(s_j, s') \cdot \mathbf{r}_1}{\Delta(s_j, s_{j+1})}\right) + \Delta(s_j, s') \cdot f_{s'}(\mathbf{r}_1)$ for some state s' and real \mathbf{r}_1 . This gives that $\text{rslope}(f_{s_j}, p_1^j - \varepsilon_j) = \text{rslope}(f_{s_{j+1}}, p_1^{j+1} - \frac{\varepsilon_j}{\Delta(s_j, s_{j+1})})$. We write $\delta = \max(\{\Delta(s, s') \mid s, s' \in S\} \setminus \{1\})$. Since $\Delta(s_j, s_{j+1}) \leq \delta$ and the slope is decreasing: $\text{rslope}(f_{s_j}, p_1^j - \varepsilon_j) \geq \text{rslope}(f_{s_{j+1}}, p_1^{j+1} - \frac{\varepsilon_j}{\delta})$.

Note that each time we repeat this, ε_j is multiplied by at least $\frac{1}{\delta}$. Hence, after finitely many steps we will reach a value greater than $\theta(s, \mathbf{p})$, which shows the property. \square

E Challenges for generalisation of our results

This section enumerates the reasons why we were unable to generalize results to multiple dimensions.

The notions of left-slope and right-slope still makes sense in three dimensions (and we could think of generalising them: front slope, back slope and other directions). However, the properties that we used in the two-dimensional case are now longer true, as we illustrate in the following lemmas. We write lslope_i for the slope in the direction of decreasing i -th dimension.

Lemma 37. *There is a bounded convex set $X \subset \mathbb{R}^3$, for which there exists two extremal points (x, y, z) and (x', y', z') such that $\text{lslope}_1(f, (x, y)) = \text{lslope}_1(f, (x', y'))$ where f_X is the function defined by $f(x, y) = \sup\{z \mid (x, y, z) \in X\}$.*

Proof. Let $X = \{(x, y, z) \mid x + y + z \leq 1 \wedge x \leq 1 \wedge y \leq 1\}$. The points $a = (1, 0, 0)$ and $b = (0, 1, 0)$ are extremal. $f_X(1 - \varepsilon, 0) = \varepsilon$ and $f_X(0 - \varepsilon, 1) = \varepsilon$ therefore $\text{lslope}_1(f_X, a) = -1 = \text{lslope}_1(f_X, b)$.

Thus to generalise Lem. 8.2 to dimension n , we would need to consider more than n directions. We could for instance consider a property of this kind:

Conjecture: If $X \subset \mathbb{R}^n$ is a bounded convex set and $p \neq p'$ are extremal points of X , then $\exists i. \text{lslope}_i(f_X, p) \neq \text{lslope}_i(f_X, p')$.

Assuming this conjecture was true, there would still be the problem of how to follow an accumulation point. In the two-dimensional case, we chose at each step an accumulation point with the same slope. Now in higher dimension, we may not be able to follow a accumulation point that has the same slope in all directions (and if the slope is not preserved in all directions, our conjecture cannot be used). We illustrate this problem, with the following lemma that shows that we could not extend the techniques used in Lem. 10.

Lemma 38. *There are bounded convex sets X, Y, Z such that $X = Y \cap Z$, $\text{lslope}_1(f_X, p) \neq \text{lslope}_1(f_Y, p)$ and $\text{lslope}_2(f_X, p) \neq \text{lslope}_2(f_Z, p)$.*

Proof. Consider $Y = \{(x, y, z) \in [0, 1]^3 \mid z \leq 1 - x\}$ and $Z = \{(x, y, z) \in [0, 1]^3 \mid z \leq 1 - y\}$ and the point $p = (\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$. Let $\varepsilon \in [0, \frac{1}{2}]$, $f_X(p - (\varepsilon, 0, 0)) = \frac{1}{2}$, so $\text{lslope}_1(f_X, p) = 0$ and similarly $\text{lslope}_2(f_X, p) = 0$. However $f_Y(p - (\varepsilon, 0, 0)) = \frac{1}{2} + \varepsilon$, so $\text{lslope}_1(f_Y, p) = -1$ and similarly $\text{lslope}_2(f_Z, p) = -1$.

This shows that the idea of following an accumulation point with the same slope cannot be generalised easily to higher number of dimensions.

This figure "drawing2.png" is available in "png" format from:

<http://arxiv.org/ps/1605.03811v1>